Working Paper
Series

FINRISK
National Centre of Competence in Research
Financial Valuation and Risk Management

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 675

Mean-Variance Hedging via Stochastic Control and BSDEs for General Semimartingales

Monique Jeanblanc

Michael Mania

Marina Santacroce

Martin Schweizer

First version: February 2011
Current version: June 2012

This research has been carried out within the NCCR FINRISK project on "Mathematical Methods in Financial Risk Management"

Mean-variance hedging via stochastic control and BSDEs for general semimartingales

Monique Jeanblanc, Université d'Evry Val d'Essonne, Laboratoire Analyse et Probabilités, IBGBI, 23 Boulevard de France, F – 91037 Evry Cedex, France and *Institut Europlace de Finance, Palais Brongniart, rue de la Bourse, F – 75002 Paris, France,* monique.jeanblanc@univ-evry.fr

Michael Mania, A. Razmadze Mathematical Institute, Aleksidze St. 1, Tbilisi 0193, and *Georgian-American University, 17a, Chavchavadze Ave, Tbilisi, Georgia,* misha.mania@gmail.com

Marina Santacroce, Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, I – 10129 Torino, Italy, marina.santacroce@polito.it

Martin Schweizer, ETH Zürich, Departement Mathematik, ETH-Zentrum, HG G 51.2, Rämistrasse 101, CH – 8092 Zürich, Switzerland,* and *Swiss Finance Institute, Walchestrasse 9, CH – 8006 Zürich, Switzerland,* martin.schweizer@math.ethz.ch

Abstract: We solve the problem of mean-variance hedging for general semimartingale models via stochastic control methods. After proving that the value process of the associated stochastic control problem has a quadratic structure, we characterise its three coefficient processes as solutions of semimartingale backward stochastic differential equations and show how they can be used to describe the optimal trading strategy for each conditional mean-variance hedging problem. For comparison with the existing literature, we provide alternative equivalent versions of the BSDEs and present a number of simple examples.

Key words: mean-variance hedging, stochastic control, backward stochastic differential equations, semimartingales, mathematical finance, variance-optimal martingale measure

MSC 2010 Classification Numbers: 60G48, 60H10, 93E20, 91G80

JEL Classification Numbers: G11, C60, C61

This version: 4.6.2012

(to appear in the Annals of Applied Probability)

* corresponding author

0. Introduction

Mean-variance hedging is one of the classical problems from mathematical finance. In financial terms, its goal is to minimise the mean squared error between a given payoff H and the final wealth of a self-financing strategy ϑ trading in the underlying assets S. Mathematically, one wants to project the random variable H in $L^2(P)$ on the space of all stochastic integrals $\vartheta \cdot S_T = \int_0^T \vartheta_r \, dS_r$, perhaps after subtracting an initial capital x. The contribution of our paper is to solve this problem via stochastic control methods and stochastic calculus techniques for the case where the asset prices S are given by a general (locally P-square-integrable) semimartingale, under a natural no-arbitrage assumption.

The literature on mean-variance hedging is vast, and we do not try to survey it here; see Schweizer (2010) for an attempt in that direction. There are two main approaches; one of them uses martingale theory and projection arguments, while the other views the task as a linear-quadratic stochastic control problem and uses backward stochastic differential equations (BSDEs) to describe the solution. By combining tools from both areas, we improve earlier work in two directions — we describe the solution more explicitly than by the martingale and projection method, and we work in a general semimartingale model without restricting ourselves to particular setups (like Itô processes or Lévy settings). We show that the value process of the stochastic control problem associated to mean-variance hedging possesses a quadratic structure, describe its three coefficient processes by semimartingale BSDEs, and show how to obtain the optimal strategy ϑ^* from there. In contrast to the majority of earlier contributions from the control strand of the literature, we also give a rigorous derivation of these BSDEs. For comparison, the usual results (especially in settings with Itô processes or jump-diffusions) start from a BSDE system and only prove a verification theorem that shows how a solution to the BSDE system induces an optimal strategy. Apart from being more precise, we think that our approach is also more informative since it shows clearly and explicitly how the BSDEs arise, and hence provides a systematic way to tackle mean-variance hedging via stochastic control in general semimartingale models. More detailed comparisons to the literature are given in the respective sections.

The paper is structured as follows. We start in Section 1 with a precise problem formulation and state the martingale optimality principle for the value process $V^H(x)$ of the associated stochastic control problem. Assuming that each (time t) conditional problem admits an optimal strategy, we then show that $V^H(x)$ is a quadratic polynomial in x whose coefficients are stochastic processes $v^{(0)}, v^{(1)}, v^{(2)}$ that do not depend on x. This is a kind of folklore result, and our only claim to originality is that we give a very simple proof in a very general setting. We also show that the coefficient $v^{(2)}$ equals the value process $V^0(1)$ for the control problem with initial value $x = 1$ and $H \equiv 0$.

Motivated by the last result, we study in Section 2 the particular problem for $x = 1$ and

$H \equiv 0$. We impose the no-arbitrage condition that there exists an equivalent σ-martingale measure for S with P-square-integrable density and are then able to characterise the process $v^{(2)}$ as the solution of a semimartingale BSDE. More precisely, Theorem 2.4 shows that all conditional problems for $x = 1, H \equiv 0$ admit optimal strategies if and only if that BSDE (2.18) has a solution in a specific class, and in that case, the unique solution is $v^{(2)}$ and the conditionally optimal strategies can be given in terms of the solution to (2.18). In comparison to earlier work, we eliminate all technical assumptions (like continuity or quasi-left-continuity) on S, and we also do not need reverse Hölder inequalities for our main results.

Section 3 considers the general case of the mean-variance hedging problem with $x \in I\!R$ and $H \in L^2(\mathcal{F}_T, P)$. The analogue of Theorem 2.4 is given in Theorem 3.1, where we describe the three coefficient processes $v^{(2)}, v^{(1)}, v^{(0)}$ by a coupled system (3.1)–(3.3) of semimartingale BSDEs. Existence of optimal strategies for all conditional problems for (x, H) is shown to be equivalent to solvability of the system (3.1)–(3.3), with solution $v^{(2)}, v^{(1)}, v^{(0)}$, and we again express the conditionally optimal strategies in terms of the solution to (3.1)–(3.3). As mentioned above, this is stronger than only a verification result.

In Section 4, we provide equivalent alternative versions for our BSDEs which are more convenient to work with in some examples with jumps. This also allows us to discuss in more detail the connections to the existing literature. Finally, Section 5 illustrates the use of our results and gives further links to the literature by a number of simple examples.

1. Problem formulation and general results

We start with a finite time horizon $T \in (0, \infty)$ and a filtered probability space $(\Omega, \mathcal{F}, I\!F, P)$ with the filtration $I\!F = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the usual conditions of right-continuity and P-completeness. Let $S = (S_t)_{0 \le t \le T}$ be an $I\!R^d$-valued RCLL semimartingale and denote by $\Theta = \Theta_S$ the space of all predictable S-integrable processes ϑ, $\vartheta \in L(S)$ for short, such that the stochastic integral process $\vartheta \cdot S = \int \vartheta \, dS$ is in the space $\mathcal{S}^2(P)$ of semimartingales. Our basic references for terminology and results from stochastic calculus are Dellacherie/Meyer (1982) and Jacod/Shiryaev (2003).

For $x \in I\!R$ and $H \in L^2(\mathcal{F}_T, P)$, the problem of *mean-variance hedging (MVH)* is to

$$\text{minimise } E\left[(H - x - \vartheta \cdot S_T)^2\right] \text{ over all } \vartheta \in \Theta. \tag{1.1}$$

The interpretation is that S models the (discounted) prices of d risky assets in a financial market containing also a riskless bank account with (discounted) price 1. An integrand ϑ together with $x \in I\!R$ then describes a self-financing dynamic trading strategy with initial wealth x, and H stands for the (discounted) payoff at time T of some financial instrument. By using (x, ϑ), we generate up to time T via trading a wealth of $x + \int_0^T \vartheta_r \, dS_r = x + \vartheta \cdot S_T$, and we want to choose ϑ in such a way that we are close, in the $L^2(P)$-sense, to the payoff H. We embed this into a *stochastic control problem* and define for $\psi \in \Theta$ and $t \in [0,T]$

$$
\begin{aligned}
V_t^H(x,\psi) &:= \operatorname*{ess\,inf}_{\vartheta\in\Theta_{t,T}(\psi)} E\left[(H - x - \vartheta\cdot S_T)^2 \,\middle|\, \mathcal{F}_t\right] \\
&= \operatorname*{ess\,inf}_{\vartheta\in\Theta_{t,T}(\psi)} E\bigg[\bigg(H - x - \int_0^t \psi_r\,dS_r - \int_t^T \vartheta_r\,dS_r\bigg)^2 \,\bigg|\, \mathcal{F}_t\bigg],
\end{aligned}
$$

where $\Theta_{t,T}(\psi) := \{\vartheta \in \Theta \,|\, \vartheta = \psi \text{ on } [\![0,t]\!]\}$. Our goal is to study the *dynamic value family*

$$
V_t^H(x) := V_t^H(x,0) = \operatorname*{ess\,inf}_{\vartheta\in\Theta} E\bigg[\bigg(H - x - \int_t^T \vartheta_r\,dS_r\bigg)^2 \,\bigg|\, \mathcal{F}_t\bigg], \qquad t \in [0,T], \tag{1.2}
$$

in order to describe the optimal strategy for the MVH problem (1.1). Observe that with these notations, we have the identity $V_u^H\Big(x + \int_t^u \psi_r\,dS_r\Big) = V_u^H(x, \psi I_{]\!]t,T]\!]}) = V_u^H(x, \psi I_{]\!]t,u]\!]})$ for $u \geq t$. Because the family of random variables $\Gamma_t(\vartheta) := E\Big[\Big(H - x - \int_t^T \vartheta_r\,dS_r\Big)^2 \,\Big|\, \mathcal{F}_t\Big]$ for $\vartheta \in \Theta$ is closed under taking maxima and minima, we have the classical *martingale optimality principle* in the following form; see for instance El Karoui (1981) for the general theory, or Mania/Tevzadze (2003a) for a formulation closer to the present one.

Proposition 1.1. *Fix $H \in L^2(\mathcal{F}_T, P)$. For every $x \in I\!R$ and $t \in [0,T]$, we have:*

1) *The process $\Big(V_u^H\Big(x + \int_t^u \vartheta_r\,dS_r\Big)\Big)_{t\leq u\leq T}$ is a P-submartingale for every $\vartheta \in \Theta$.*

2) *A strategy $\vartheta^{*,t} = \vartheta^{*,t}(x,H) \in \Theta_{t,T}(0)$ is optimal for (1.2) (i.e. attains the essential infimum there) if and only if $\Big(V_u^H\Big(x + \int_t^u \vartheta_r^{*,t}\,dS_r\Big)\Big)_{t\leq u\leq T}$ is a P-martingale.*

3) *If $\vartheta^* = \vartheta^{*,0}(x,H)$ solves (1.1), $\vartheta^* I_{]\!]t,T]\!]}$ is optimal for $V_t^H(x + \vartheta^*\cdot S_t) = V_t^H(x,\vartheta^*)$.*

For the special case $H \equiv 0$, the fact that Θ is a cone immediately gives

$$
V_t^0(x) = \operatorname*{ess\,inf}_{\vartheta\in\Theta} E\bigg[\bigg(x + \int_t^T \vartheta_r\,dS_r\bigg)^2 \,\bigg|\, \mathcal{F}_t\bigg] = x^2 V_t^0(1). \tag{1.3}
$$

This holds for any random variable $x \in L^2(\mathcal{F}_t, P)$. So Proposition 1.1 almost directly gives

Corollary 1.2. *For every $t \in [0,T]$, we have:*

1) *The process $\Big(\Big(1 + \int_t^u \vartheta_r\,dS_r\Big)^2 V_u^0(1)\Big)_{t\leq u\leq T}$ is a P-submartingale for every $\vartheta \in \Theta$.*

2) *A strategy $\vartheta^{*,t} = \vartheta^{*,t}(1,0) \in \Theta_{t,T}(0)$ is optimal for $V_t^0(1)$ in (1.3) if and only if the process $\Big(\Big(1 + \int_t^u \vartheta_r^{*,t}\,dS_r\Big)^2 V_u^0(1)\Big)_{t\leq u\leq T}$ is a P-martingale.*

3) *If $\vartheta^* = \vartheta^{*,0}(1,0)$ solves (1.1) for $x = 1$ and $H \equiv 0$, then*

$$(1.4) \qquad \int_t^T \vartheta^*_r \, dS_r = 0 \qquad P\text{-a.s. on the set } \{1 + \vartheta^* \cdot S_t = 0\}.$$

Proof. Since 1) and 2) are special cases of Proposition 1.1, we only need to prove 3). Fix $t \in [0,T]$, set $D_t := \{1 + \vartheta^* \cdot S_t = 0\} \in \mathcal{F}_t$ and define $\varphi := I_{D_t^c} \vartheta^* I_{]\!]t,T]\!]}$. By part 3) of Proposition 1.1 with $x = 1, H \equiv 0$, the strategy $\vartheta^* I_{]\!]t,T]\!]}$ is optimal for $V_t^0(1 + \vartheta^* \cdot S_t)$ so that

$$I_{D_t} E\bigg[\Big(1 + \vartheta^* \cdot S_t + \int_t^T \vartheta^*_r \, dS_r\Big)^2 \,\bigg|\, \mathcal{F}_t\bigg] \le I_{D_t} E\bigg[\Big(1 + \vartheta^* \cdot S_t + \int_t^T \varphi_r \, dS_r\Big)^2 \,\bigg|\, \mathcal{F}_t\bigg] = 0$$

by the definitions of φ and D_t. This yields $0 = I_{D_t}\Big(1 + \vartheta^* \cdot S_t + \int_t^T \vartheta^*_r \, dS_r\Big) = I_{D_t} \int_t^T \vartheta^*_r \, dS_r$ P-a.s. again by the definition of D_t, and so we get (1.4). **q.e.d.**

As in Proposition A.2 of Mania/Tevzadze (2003a) or Théorème 2.28 of El Karoui (1981), we also obtain

Proposition 1.3. *Fix $H \in L^2(\mathcal{F}_T, P)$. For every $x \in I\!R$, $t \in [0,T]$ and $\psi \in \Theta$, there exists an RCLL version of the P-submartingale $\Big(V_u^H\Big(x + \int_t^u \psi_r \, dS_r\Big)\Big)_{t \le u \le T}$. Moreover, for each $x \in I\!R$, the family $\{V_t^H(x) \,|\, t \in [0,T]\}$ of random variables can be aggregated into an RCLL process, which we again call $V^H(x) = (V_u^H(x))_{0 \le u \le T}$.*

In the sequel, we always choose and work with the RCLL versions from Proposition 1.3.

For easier discussion of the next result, we introduce some more terminology. We denote by $I\!P^2_{\mathrm{e},\sigma}(S)$ the (a priori possibly empty) set of all probability measures Q equivalent to P on $\mathcal{F}_T$ such that S is a Q-σ-martingale and $\frac{dQ}{dP} \in L^2(P)$. Assuming that $I\!P^2_{\mathrm{e},\sigma}(S)$ is nonempty is one way of imposing *absence of arbitrage* for our financial market and also fits naturally with the fact that our basic problem is cast in quadratic terms. The density process of Q with respect to P is denoted by $Z^Q = (Z_t^Q)_{0 \le t \le T}$, and we say that $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ satisfies the reverse Hölder inequality $R_2(P)$ if there is a constant C with $E_P[(Z_T^Q)^2 \,|\, \mathcal{F}_\tau] \le C(Z_\tau^Q)^2$ P-a.s. for all stopping times $\tau \le T$. It is well known that if there is some $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ satisfying $R_2(P)$, then $G_T(\Theta) = \{\vartheta \cdot S_T \,|\, \vartheta \in \Theta\}$ as well as $L^2(\mathcal{F}_t, P) + G_T(\Theta_{t,T}(0))$ for each t are closed in $L^2(P)$ so that both (1.1) and (1.2) for each t have a solution; see Theorem 5.2 of Choulli/Krawczyk/Stricker (1998). Moreover, for any $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ and any $\vartheta \in \Theta$, the product of Z^Q and $\vartheta \cdot S$ is a P-σ-martingale with P-integrable supremum; so $\vartheta \cdot S$ is a true Q-martingale, and $\vartheta \cdot S_T = 0$ a.s. implies that $\vartheta = 0$ in $L(S)$. This is used later several times to argue that a self-financing strategy is uniquely determined by its wealth process (i.e. stochastic integral).

Our main result in this section now provides the basic structure of the process $V^H(x)$ and of the optimal strategies for (1.2).

Theorem 1.4. *Fix $H \in L^2(\mathcal{F}_T, P)$. Suppose that for each $t \in [0,T]$, (1.2) has a solution $\vartheta^{*,t} = \vartheta^{*,t}(x,H)$ for every $x \in \mathbb{R}$. Suppose also that for any $\vartheta \in \Theta$, $\vartheta \cdot S_T = 0$ a.s. implies that $\vartheta = 0$ in $L(S)$. Then each $\vartheta^{*,t}(x,H)$ is of the affine form*

$$\vartheta^{*,t}(x,H) = \vartheta^{0,t} + x\vartheta^{1,t} \qquad \text{for some } \vartheta^{0,t}, \vartheta^{1,t} \in \Theta_{t,T}(0) \tag{1.5}$$

and each $V_t^H(x)$ has the quadratic form

$$V_t^H(x) = v_t^{(0)} - 2v_t^{(1)}x + v_t^{(2)}x^2 \tag{1.6}$$

for RCLL processes $v^{(0)}, v^{(1)}, v^{(2)}$ not depending on x. Moreover, $\vartheta^{1,t} = \vartheta^{*,t}(1,0)$ is the solution of (1.3), and the quadratic coefficient $v_t^{(2)}$ equals $V_t^0(1)$ from (1.3) and does not depend on H.

Proof. Fix $t \in [0,T]$. Denote by $G_{t,T} = G_T(\Theta_{t,T}(0)) = \left\{ \int_t^T \vartheta_r \, dS_r \,\middle|\, \vartheta \in \Theta \right\}$ the space of all stochastic integrals on $]\!]t,T]\!]$ of $\vartheta \in \Theta$ and by $\overline{G_{t,T}}$ its closure in $L^2(P)$. Since the problems (1.2) with payoff H for $x = 1$ and $x = 0$ have solutions (which are given by projections), so does the problem (1.2) for $x = 1$ and payoff $H' \equiv 0$ by taking differences, and the latter problem is identical to (1.2) for $x = 0, H' \equiv -1$ so that $\vartheta^*(0,-1) = \vartheta^*(1,0)$. Both here and in the next argument, we exploit our assumption that a self-financing strategy is uniquely determined by its wealth process. If Π is the projection in $L^2(P)$ on $\overline{G_{t,T}}$, then clearly

$$\vartheta^{*,t}(x,H) \cdot S_T = \Pi(H - x) = \Pi(H) + x\Pi(-1) = \vartheta^{*,t}(0,H) \cdot S_T + x\vartheta^{*,t}(0,-1) \cdot S_T,$$

and so (1.5) follows with $\vartheta^{0,t} = \vartheta^{*,t}(0,H)$ and $\vartheta^{1,t} = \vartheta^{*,t}(0,-1) = \vartheta^{*,t}(1,0)$. This gives

$$V_t^H(x) = E\left[\left(H - x - \int_t^T \vartheta_r^{*,t}(x,H)\, dS_r\right)^2 \,\middle|\, \mathcal{F}_t\right] = E\left[\left(H - \int_t^T \vartheta_r^{0,t}\, dS_r - x\left(1 + \int_t^T \vartheta_r^{1,t}\, dS_r\right)\right)^2 \,\middle|\, \mathcal{F}_t\right],$$

and hence we directly obtain the expression (1.6) with $v_t^{(0)} = E\left[\left(H - \int_t^T \vartheta_r^{*,t}(0,H)\, dS_r\right)^2 \,\middle|\, \mathcal{F}_t\right]$, $v_t^{(1)} = E\left[\left(H - \int_t^T \vartheta_r^{*,t}(0,H)\, dS_r\right)\left(1 + \int_t^T \vartheta_r^{*,t}(1,0)\, dS_r\right) \,\middle|\, \mathcal{F}_t\right]$ and

$$v_t^{(2)} = E\left[\left(1 + \int_t^T \vartheta_r^{*,t}(1,0)\, dS_r\right)^2 \,\middle|\, \mathcal{F}_t\right] = V_t^0(1). \tag{1.7}$$

Since the families $\{V_t^H(x) \,|\, t \in [0,T]\}$ aggregate into an RCLL process, the same holds for the families $v^{(0)}, v^{(1)}, v^{(2)}$ from (1.6). The last assertion is clear from the above proof. **q.e.d.**

Remarks. 1) As mentioned above, one sufficient condition for all assumptions of Theorem 1.4 is the existence of some $Q \in I\!\!P^2_{e,\sigma}(S)$ satisfying the reverse Hölder inequality $R_2(P)$; see Choulli/Krawczyk/Stricker (1998).

2) The particular choice of $\Theta = \Theta_S$ for the space of integrands is convenient and also exploited later, but not crucially important for the conclusion of Theorem 1.4 to hold. All we need is that there exist for all t solutions $\vartheta^{*,t}(x,H)$ for all x, that the martingale optimality principle from Proposition 1.1 holds, and that Θ (or $G_T(\Theta)$, which must be a subset of $L^2(P)$) is a linear space. Of course, existence of solutions for all x and all H is equivalent to closedness of $G_T(\Theta)$ in $L^2(P)$; and the key point for the martingale optimality principle is closedness under bifurcation of Θ.

3) We emphasise that Theorem 1.4 is a bit of a folklore result in the literature on mean-variance hedging, and we do not claim any great originality here. Variants in different levels of generality can be found in Gugushvili (2003), Mania/Tevzadze (2003a), Bobrovnytska/Schweizer (2004), Černý (2004), to name but a few. However, we think that it is useful to have a presentation which is as general and yet as simple as possible. ◇

Our goal in the sequel is to study the dynamics of the coefficient processes $v^{(0)}, v^{(1)}, v^{(2)}$ and use them to express the optimal strategies $\vartheta^{*,t}(x,H)$. Let us first simplify things a little. Because $\vartheta^{*,t}(1,0)$ is the solution (minimiser) of (1.3), the first order condition for that quadratic optimisation problem implies that $E\Big[\int_t^T \vartheta_r \, dS_r \Big(1 + \int_t^T \vartheta^{*,t}_r(1,0)\, dS_r\Big) \,\Big|\, \mathcal{F}_t\Big] = 0$ P-a.s. for each $t \in [0,T]$ and $\vartheta \in \Theta$. We note for later use that this allows us to write

$$v_t^{(1)} = E\Big[H\Big(1 + \int_t^T \vartheta^{*,t}_r(1,0)\, dS_r\Big) \,\Big|\, \mathcal{F}_t\Big]. \tag{1.8}$$

Also for later use, we give some additional results for the coefficients $v^{(0)}, v^{(1)}, v^{(2)}$.

Lemma 1.5 *Under the assumptions of Theorem 1.4, we have:*

1) $v^{(2)}$ *is a P-submartingale with* $0 \le v^{(2)} \le 1$.

2) $v^{(0)}$ *is a P-submartingale with* $0 \le v_t^{(0)} \le E[H^2 \,|\, \mathcal{F}_t]$, $0 \le t \le T$, *hence of class (D).*

3) $v^{(1)}$ *is a P-special semimartingale with $|v^{(1)}|^2$ of class (D). Therefore $v^{(1)}$ is in $\mathcal{S}^2_{\mathrm{loc}}(P)$ and for its canonical decomposition $v^{(1)} = v_0^{(1)} + m^{(1)} + a^{(1)}$, we have $m^{(1)} \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$.*

Proof. 1) By Theorem 1.4 and (1.7), we have $v^{(2)} = V^0(1)$, and this is a P-submartingale by part 1) of Corollary 1.2 (for $\vartheta \equiv 0$). Because $\vartheta \equiv 0$ is in Θ, we get $0 \le V^0(1) \le 1$ directly from (1.3).

2) Theorem 1.4 gives $v^{(0)} = V^H(0)$, and this is a P-submartingale by part 1) of Proposition 1.1 (for $x = 0, \vartheta \equiv 0$) and nonnegative by the definition in (1.2). Since $\vartheta \equiv 0$ is in Θ, (1.2) also gives $V_t^H(0) \leq E[H^2 \,|\, \mathcal{F}_t]$ for all t.

3) By part 1) of Proposition 1.1, $V^H(x)$ is a P-submartingale, hence a P-special semimartingale, and so are $v^{(2)}$ and $v^{(0)}$ by 1) and 2). Because $V^H(x) = v^{(0)} - 2v^{(1)}x + v^{(2)}x^2$ by Theorem 1.4, also $v^{(1)}$ is then a P-special semimartingale. Moreover, $V^H(x) \geq 0$ for all x due to (1.2) implies that $|v_t^{(1)}|^2 \leq v_t^{(2)} v_t^{(0)} \leq v_t^{(0)} \leq E[H^2 \,|\, \mathcal{F}_t]$, $0 \leq t \leq T$, by 1) and 2) so that $|v^{(1)}|^2$ is of class (D). The rest of part 3) is then clear. **q.e.d.**

2. Pure investment: the special case $x = 1$, $H \equiv 0$

In this section, we give a description of (the RCLL version of) the value process

$$V_t^0(1) = \operatorname*{ess\,inf}_{\vartheta \in \Theta} E\bigg[\Big(1 + \int_t^T \vartheta_r \, dS_r\Big)^2 \,\bigg|\, \mathcal{F}_t\bigg], \qquad 0 \leq t \leq T \tag{2.1}$$

of the problem (1.3). Since this is by (1.7) and Theorem 1.4 the *quadratic* coefficient in the representation (1.6), we use in this section the shorter notation

$$q_t := V_t^0(1) = v_t^{(2)}, \qquad 0 \leq t \leq T.$$

We also remark that q coincides with the *opportunity process* from Černý/Kallsen (2007), although the latter is defined there with a different space Θ of integrands ϑ for S.

Let us first prove strict positivity of q, as well as of q_-.

Lemma 2.1. *Suppose $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$. Then q and q_- are both strictly positive, in the sense that $P[q_t > 0$ and $q_{t-} > 0$ for $0 \leq t \leq T] = 1$. If there is some $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ satisfying the reverse Hölder inequality $R_2(P)$, we even have $q \geq \delta > 0$ P-a.s. for some constant δ.*

Proof. For $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ with density process $Z = Z^Q = Z^{Q;P}$, define as in Gouriéroux/Laurent/Pham (1998) a new probability $R \approx P$ by $\frac{dR}{dP} := \frac{Z_T^2}{E[Z_T^2]}$. Then the Bayes rule gives

$$Z_t^{R;P} := \frac{dR}{dP}\Big|_{\mathcal{F}_t} = \frac{E[Z_T^2 \,|\, \mathcal{F}_t]}{E[Z_T^2]}, \tag{2.2}$$

$$Z_t^{R;Q} := \frac{dR}{dQ}\Big|_{\mathcal{F}_t} = \frac{E_Q[Z_T \,|\, \mathcal{F}_t]}{E[Z_T^2]} = \frac{1}{Z_t} Z_t^{R;P}. \tag{2.3}$$

Using the Bayes rule and (2.2), Jensen's inequality, again the Bayes rule and (2.3) yields

$$
\begin{aligned}
E\Big[\Big(1+\int_t^T \vartheta_r\,dS_r\Big)^2\,\Big|\,\mathcal{F}_t\Big] &= Z_t^{R;P} E[Z_T^2] E_R\Big[(Z_T^2)^{-1}\Big(1+\int_t^T \vartheta_r\,dS_r\Big)^2\,\Big|\,\mathcal{F}_t\Big] \\
&\geq Z_t^{R;P} E[Z_T^2]\Big(E_R\Big[(Z_T)^{-1}\Big(1+\int_t^T \vartheta_r\,dS_r\Big)\,\Big|\,\mathcal{F}_t\Big]\Big)^2 \\
&= Z_t^{R;P} E[Z_T^2]\Big((Z_t^{R;Q})^{-1} E_Q\Big[(E[Z_T^2])^{-1}\Big(1+\int_t^T \vartheta_r\,dS_r\Big)\,\Big|\,\mathcal{F}_t\Big]\Big)^2.
\end{aligned}
$$

But as already noted before Theorem 1.4, $\int \vartheta\,dS$ is a Q-martingale whenever $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ and $\vartheta \in \Theta$. So we get by using (2.3) and (2.2) that

$$
E\Big[\Big(1+\int_t^T \vartheta_r\,dS_r\Big)^2\,\Big|\,\mathcal{F}_t\Big] \geq \frac{Z_t^{R;P} E[Z_T^2]}{(Z_t^{R;Q} E[Z_T^2])^2} = \frac{Z_t^2}{E[Z_T^2\,|\,\mathcal{F}_t]}, \tag{2.4}
$$

and the first assertion follows since $\inf_{0\leq t\leq T} Z_t > 0$ P-a.s. by the minimum principle for supermartingales and $\sup_{0\leq t\leq T} E[Z_T^2\,|\,\mathcal{F}_t] < \infty$ P-a.s. by the martingale maximal inequality. If Q satisfies $R_2(P)$ with constant C, we can take $\delta = 1/C$ for the second claim. **q.e.d.**

Remark. Strict positivity of the opportunity process and its left limits (hence of q and q_-) is also proved in Lemma 3.10 of Černý/Kallsen (2007). However, the above short proof seems to us more transparent. ◇

The optimisation problem in (2.1) has a (well-known) dual formulation as follows. Extending $I\!P^2_{\mathrm{e},\sigma}(S)$ a little, we denote by $I\!P^2_{\mathrm{s},\sigma}(S)$ the set of all *signed* measures $Q \ll P$ on $\mathcal{F}_T$ with $Q[\Omega] = 1$ and such that the product of S and the density process Z^Q of Q with respect to P is a P-σ-martingale. We call $\widetilde{Q} \in I\!P^2_{\mathrm{s},\sigma}(S)$ *variance-optimal* if $\|\frac{d\widetilde{Q}}{dP}\|_{L^2(P)} \leq \|\frac{dQ}{dP}\|_{L^2(P)}$ for all $Q \in I\!P^2_{\mathrm{s},\sigma}(S)$, and we say that the *variance-optimal martingale measure (VOMM)* exists if $\widetilde{Q} \in I\!P^2_{\mathrm{e},\sigma}(S)$ is variance-optimal. (In particular, $\widetilde{Q}$ is then by definition equivalent to P.) If S is continuous, Theorem 1.3 of Delbaen/Schachermayer (1996) shows that $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$ is sufficient for the VOMM to exist; but if S can have jumps, the situation is more complicated.

The dynamic problem of finding the VOMM has the value process

$$
\widetilde{V}_t := \operatorname*{ess\,inf}_{Q\in I\!P^2_{\mathrm{e},\sigma}(S)} E\big[\big(Z_T^Q/Z_t^Q\big)^2\,\big|\,\mathcal{F}_t\big], \qquad 0 \leq t \leq T.
$$

Then we have the following direct connection to $V^0(1)$ and (2.1).

Proposition 2.2. *Suppose $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and that the VOMM exists. Then $\widetilde{V} = 1/V^0(1)$.*

Proof. We know from (2.4) in the proof of Lemma 2.1 that for $\vartheta \in \Theta$ and $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$,

$$E\Big[\Big(1+\int_t^T \vartheta_r\,dS_r\Big)^2\,\Big|\,\mathcal{F}_t\Big] \ge 1/E\big[\big(Z^Q_T/Z^Q_t\big)^2\,\big|\,\mathcal{F}_t\big], \qquad 0 \le t \le T.$$

Taking the ess inf over $\vartheta \in \Theta$ and the ess sup over $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ gives $V^0(1) \ge 1/\widetilde{V}$. Conversely, since $V^0_T(1) = 1$, the martingale optimality principle in Corollary 1.2 gives

$$\Big(1+\int_0^t \vartheta_r\,dS_r\Big)^2 V^0_t(1) \le E\Big[\Big(1+\int_0^T \vartheta_r\,dS_r\Big)^2\,\Big|\,\mathcal{F}_t\Big], \qquad 0 \le t \le T \tag{2.5}$$

for every $\vartheta \in \Theta = \Theta_S$. But if we define, as in Gouriéroux/Laurent/Pham (1998),

$$\Theta_{\mathrm{GLP}} := \{\vartheta \in L(S)\,|\,\vartheta\cdot S_T \in L^2(P) \text{ and } Z^Q(\vartheta\cdot S) \text{ is a } P\text{-martingale for all } Q \in I\!P^2_{\mathrm{s},\sigma}(S)\},$$

then Corollary 2.9 of Černý/Kallsen (2007) says that $G_T(\Theta_{\mathrm{GLP}}) := \{\vartheta\cdot S_T\,|\,\vartheta \in \Theta_{\mathrm{GLP}}\}$ is the closure of $G_T(\Theta_S)$ in $L^2(P)$, and this allows us to extend (2.5) to every $\vartheta \in \Theta_{\mathrm{GLP}}$. Indeed, for a sequence (ϑ^n) in Θ_S with $G_T(\vartheta^n) \to G_T(\vartheta)$ in $L^2(P)$, the right-hand side of (2.5) for ϑ^n converges in $L^1(P)$ to the right-hand side of (2.5) for ϑ, and because we have $Z^Q_t\Big(1+\int_0^t \vartheta_r\,dS_r\Big) = E\Big[Z^Q_T\Big(1+\int_0^T \vartheta_r\,dS_r\Big)\,\Big|\,\mathcal{F}_t\Big]$ for $\vartheta \in \Theta_{\mathrm{GLP}} \supseteq \Theta_S$ and $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$, the left-hand side of (2.5) for ϑ^n converges in probability to the left-hand side of (2.5) for ϑ. We remark that the use of Corollary 2.9 in Černý/Kallsen (2007) exploits that $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$.

By assumption, the VOMM $\widetilde{Q}$ exists. A slight modification of the proof of Lemma 2.2 in Delbaen/Schachermayer (1996) (since S is in $\mathcal{S}^2_{\mathrm{loc}}(P)$ instead of locally bounded) yields $Z^{\widetilde{Q}}_T = c + \int_0^T \widetilde{\vartheta}_r\,dS_r$ for some $c > 0$ and $\widetilde{\vartheta} \in \Theta_{\mathrm{GLP}}$ and thus $E_{\widetilde{Q}}[Z^{\widetilde{Q}}_T\,|\,\mathcal{F}_t] = c + \int_0^t \widetilde{\vartheta}_r\,dS_r$, $0 \le t \le T$. Applying (2.5) with $\vartheta := \widetilde{\vartheta}/c$ and using the Bayes rule therefore gives

$$(Z^{\widetilde{Q}}_t)^2 E\big[\big(Z^{\widetilde{Q}}_T\big)^2\,\big|\,\mathcal{F}_t\big] \ge (Z^{\widetilde{Q}}_t)^2\big(E_{\widetilde{Q}}[Z^{\widetilde{Q}}_T\,|\,\mathcal{F}_t]\big)^2 V^0_t(1) = \big(E\big[\big(Z^{\widetilde{Q}}_T\big)^2\,\big|\,\mathcal{F}_t\big]\big)^2 V^0_t(1)$$

and hence

$$1/V^0_t(1) \ge E\big[\big(Z^{\widetilde{Q}}_T/Z^{\widetilde{Q}}_t\big)^2\,\big|\,\mathcal{F}_t\big] \ge \widetilde{V}_t, \qquad 0 \le t \le T.$$

This completes the proof. **q.e.d.**

Remark. For experts on mean-variance hedging, Proposition 2.2 is also a kind of folklore result. For the case where the filtration is continuous, it can for instance be found in Proposition 4.2 of Mania/Tevzadze (2003a) (with the remark that it extends to general $I\!F$ if S is continuous). But we do not know a reference for the level of generality given here. ◇

We often use below the following simple fact:

(2.6) If B, C are of locally integrable variation and $B \ll C$, then also $B^{\mathbf{p}} \ll C^{\mathbf{p}}$.

In (2.6), the (right) superscript $^{\mathbf{p}}$ denotes the compensator or dual predictable projection. This should not be confused with the predictable projection of a process Y which is denoted by ${}^{\mathbf{p}}Y$, with a left superscript. The most frequent application of (2.6) will be for $C = [M]$, where $C^{\mathbf{p}} = [M]^{\mathbf{p}} = \langle M \rangle$ when M is a locally square-integrable local martingale.

In the sequel, we focus on the case $d = 1$ so that S is one-dimensional. One can obtain analogous results for $d > 1$ (and we shall comment on this later), but the arguments and formulations look more technical without providing extra insight. When $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ so that S is in particular a P-special semimartingale, we write $S = S_0 + M + A$ for its P-canonical decomposition and note that $M \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ and A is predictable and of locally square-integrable (or even locally bounded) variation. If we also have $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$, then it is well known that S satisfies the so-called structure condition, i.e. that S has the form

$$S = S_0 + M + A = S_0 + M + \int \lambda \, d\langle M \rangle \tag{2.7}$$

with $M \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ and $\lambda \in L^2_{\mathrm{loc}}(M)$; see Theorem 1 of Schweizer (1995). This implies that

$$[A] = \left[\int \lambda \, d\langle M \rangle \right] = \sum (\lambda_s \Delta \langle M \rangle_s)^2 = (\lambda^2 \Delta \langle M \rangle) \boldsymbol{\cdot} \langle M \rangle \ll \langle M \rangle.$$

Because A is predictable, $[M, A]$ is a local P-martingale by Yoeurp's lemma so that

$$[S]^{\mathbf{p}} = ([M] + [A])^{\mathbf{p}} = (1 + \lambda^2 \Delta \langle M \rangle) \boldsymbol{\cdot} \langle M \rangle. \tag{2.8}$$

Now suppose that $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$. To describe the process $q = V^0(1)$ by a BSDE, we first introduce an auxiliary operation. Suppose Y is a P-special semimartingale with canonical decomposition $Y = Y_0 + N^Y + B^Y$. Then $[Y, [S]] = [N^Y, [S]] + \Delta B^Y \boldsymbol{\cdot} [S]$, and if $[N^Y, [S]]$ is of locally P-integrable variation, we have by (2.8) and (2.6) that

$$\big[Y, [S]\big]^{\mathbf{p}} = \big[N^Y, [S]\big]^{\mathbf{p}} + \Delta B^Y \boldsymbol{\cdot} [S]^{\mathbf{p}} \ll \langle M \rangle. \tag{2.9}$$

Note also that the predictable stopping theorem gives $\Delta B^Y = {}^{\mathbf{p}}\Delta Y = {}^{\mathbf{p}}Y - Y_-$ so that

$$Y_- + \Delta B^Y = {}^{\mathbf{p}}Y. \tag{2.10}$$

The auxiliary quantity we need is the predictable Radon–Nikodým derivative

$$g_t(Y) := \frac{d\big[N^Y, [S]\big]^{\mathbf{p}}_t}{d\langle M \rangle_t}, \qquad 0 \le t \le T. \tag{2.11}$$

Finally, we introduce the notation

$$\mathcal{N}(Y) := {}^{\mathbf{p}}Y \,(1 + \lambda^2 \Delta \langle M \rangle) + g(Y). \tag{2.12}$$

The condition that $[N^Y, [S]\,]$ is in $\mathcal{A}_{\mathrm{loc}}(P)$ (and hence has a compensator) is for instance satisfied if Y is bounded, hence in particular for $Y = q$.

Remark. In the context of the equations we study, the operation $\mathcal{N}(Y)$ in (2.12) can sometimes be simplified. If S is continuous, then so are $[S]$ and $\langle M\rangle$, due to (2.7); so $g(Y)$ and $\Delta\langle M\rangle$ then both vanish and (2.12) reduces to $\mathcal{N}(Y) = {}^{\mathbf{p}}Y = Y_- + \Delta B^Y$. Looking ahead at (2.18), however, we see that we are interested in the case where $B^Y \ll \langle M\rangle$, and so we then also get $\Delta B^Y = 0$ and hence $N^Y = Y_-$. Finally, if even the filtration $I\!\!F$ is continuous, then L in (2.18) is continuous; so is then Y, and we end up with $\mathcal{N}(Y) = Y$. ◇

Our next result shows that $\mathcal{N}(q) = \mathcal{N}(v^{(2)})$ is always strictly positive. This is important since we later need to divide by $\mathcal{N}(q)$.

Lemma 2.3. *Suppose $I\!\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$ and $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$. If $q \geq \delta > 0$ for some constant δ, then*

$$\mathcal{N}(q) = {}^{\mathbf{p}}q\,(1 + \lambda^2\Delta\langle M\rangle) + g(q) \geq \delta \qquad P\otimes\langle M\rangle\text{-a.e. on } [\![0,T]\!]. \tag{2.13}$$

In general, we still have

$$\mathcal{N}(q) > 0 \qquad P\otimes\langle M\rangle\text{-a.e. on } [\![0,T]\!]. \tag{2.14}$$

Moreover, $\mathcal{N}(q)$ is locally bounded away from 0 (uniformly in t,ω).

Proof. If $q \geq \delta$, then $B := q\boldsymbol{\cdot}[S] - \delta[S]$ is in $\mathcal{A}^+_{\mathrm{loc}}(P)$ and hence also $B^{\mathbf{p}} \in \mathcal{A}^+_{\mathrm{loc}}(P)$. But $B \ll [S]$, hence $B^{\mathbf{p}} \ll [S]^{\mathbf{p}} = (1 + \lambda^2\Delta\langle M\rangle)\boldsymbol{\cdot}\langle M\rangle$ by (2.6) and (2.8), and so

$$B^{\mathbf{p}} = (q\boldsymbol{\cdot}[S])^{\mathbf{p}} - \delta(1+\lambda^2\Delta\langle M\rangle)\boldsymbol{\cdot}\langle M\rangle = \int \Big(\tfrac{d(q\boldsymbol{\cdot}[S])^{\mathbf{p}}}{d\langle M\rangle} - \delta(1+\lambda^2\Delta\langle M\rangle)\Big)\,d\langle M\rangle \in \mathcal{A}^+_{\mathrm{loc}}(P).$$

Writing $q = q_- + \Delta q$ and $\Delta q\boldsymbol{\cdot}[S] = [q,[S]\,]$ and using (2.8)–(2.12) yields

$$\begin{aligned}(q\boldsymbol{\cdot}[S])^{\mathbf{p}} &= q_-\boldsymbol{\cdot}[S]^{\mathbf{p}} + \big[N^q,[S]\big]^{\mathbf{p}} + \Delta B^q\boldsymbol{\cdot}[S]^{\mathbf{p}}\\ &= \big({}^{\mathbf{p}}q\,(1+\lambda^2\Delta\langle M\rangle) + g(q)\big)\boldsymbol{\cdot}\langle M\rangle\\ &= \mathcal{N}(q)\boldsymbol{\cdot}\langle M\rangle.\end{aligned} \tag{2.15}$$

Thus we obtain $B^{\mathbf{p}} = \{\mathcal{N}(q) - \delta(1+\lambda^2\Delta\langle M\rangle)\}\boldsymbol{\cdot}\langle M\rangle \in \mathcal{A}^+_{\mathrm{loc}}(P)$, and this implies (2.13) since $\lambda^2\Delta\langle M\rangle \geq 0$. In general, setting $\tau_n := \inf\{t \geq 0 \,|\, q_t < \frac{1}{n}\} \wedge T$ (with $\inf\emptyset = +\infty$) gives $\tau_n \nearrow T$ stationarily because $q > 0$ by Lemma 2.1, and $q \geq \frac{1}{n}$ on $D_n := [\![0,\tau_n[\![\cup (\Omega\times\{T\})$ since $q_T = 1$. The argument for (2.13) now implies that $\mathcal{N}(q) \geq \frac{1}{n}$ holds $P\otimes\langle M\rangle$-a.e. on D_n, and (2.14) follows since $\bigcup_{n\in I\!\!N} D_n = [\![0,T]\!]$. For the final assertion, note that the preceding proof shows that $\mathcal{N}(q)^{\tau_n -} \geq \frac{1}{n}$ so that the nonnegative process $1/\mathcal{N}(q)$ is prelocally bounded. Since $1/\mathcal{N}(q)$ is like $\mathcal{N}(q)$ predictable, it is therefore by Dellacherie/Meyer (1982), VIII.11 also locally bounded, and this means that $\mathcal{N}(q)$ is locally bounded away from 0. **q.e.d.**

Remark. If $d > 1$, both $[S]$ and $\langle M \rangle$ have to be replaced by matrix-valued processes $([S^i, S^j])_{i,j=1,\dots,d}$ and $(\langle M^i, M^j \rangle)_{i,j=1,\dots,d}$. We then take a predictable $B \in \mathcal{A}^+_{\mathrm{loc}}(P)$ with $\langle M^i, M^j \rangle = \mu^{ij} \cdot B \ll B$ and define the matrix-valued predictable process $g(q)$ by

$$g^{ij}_t(q) := \frac{d\big[N^q, [S^i, S^j]\big]^{\mathbf{p}}_t}{dB_t}, \qquad 0 \leq t \leq T. \tag{2.16}$$

Analogously to Lemma 2.3, one can then prove that

$$\mathcal{N}(q) := {}^{\mathbf{p}}q\,\big(\mu + (\mu\lambda)^{\mathrm{tr}}\mu\lambda\Delta B\big) + g(q) \text{ is positive definite } P \otimes B\text{-a.e.} \tag{2.17}$$

◇

Recalling the notation (2.12), we now consider the *backward equation*

$$\begin{aligned} Y_t &= Y_0 + \int_0^t \frac{(\psi_s + \lambda_s {}^{\mathbf{p}}Y_s)^2}{\mathcal{N}_s(Y)}\, d\langle M \rangle_s + \int_0^t \psi_s\, dM_s + L_t \\ &= Y_0 + \int_0^t \frac{(\psi_s + \lambda_s {}^{\mathbf{p}}Y_s)^2}{{}^{\mathbf{p}}Y_s(1 + \lambda_s^2 \Delta\langle M \rangle_s) + g_s(Y)}\, d\langle M \rangle_s + \int_0^t \psi_s\, dM_s + L_t, \qquad Y_T = 1. \end{aligned} \tag{2.18}$$

A *solution* of (2.18) is a triple (Y, ψ, L), where L is a local P-martingale which is strongly P-orthogonal to M, ψ is in $L^1_{\mathrm{loc}}(M)$, and $Y = Y_0 + N^Y + B^Y$ is a P-special semimartingale with $[N^Y, [S]] \in \mathcal{A}_{\mathrm{loc}}(P)$. Note that λ and M come from S via (2.7). With a slight abuse of terminology, we sometimes call Y instead of the whole triple (Y, ψ, L) a solution; any properties then only refer to Y.

Denoting the stochastic exponential started at time t of a semimartingale X by

$${}^t\mathcal{E}(X)_u = 1 + \int_t^u {}^t\mathcal{E}(X)_{r-}\, dX_r = \mathcal{E}(X - X^t)_u, \qquad t \leq u \leq T,$$

our first main result is the following description of $V^0(1) = q$ via a BSDE.

Theorem 2.4. *Suppose that $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $\mathbb{P}^2_{\mathrm{e},\sigma}(S) \neq \emptyset$. Then:*

1) *The following two assertions are equivalent:*

a) *For every $t \in [0, T]$, there exists an optimal strategy $\vartheta^{*,t}(1, 0) \in \Theta_{t,T}(0)$ for (1.2) with $x = 1, H \equiv 0$.*

b) *There exists a solution (Y, ψ, L) to the BSDE (2.18) having $L \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$, $\psi \in L^2_{\mathrm{loc}}(M)$, Y bounded and strictly positive, and such that for every $t \in [0, T]$, the process $({}^t\mathcal{E}(-\frac{\psi + \lambda {}^{\mathbf{p}}Y}{\mathcal{N}(Y)} \cdot S)_u)_{t \leq u \leq T}$ is in $\mathcal{S}^2(P)$.*

If a) or b) hold, then the optimal $\vartheta^{*,t}(1,0)$ is for every t given by

$$\vartheta^{*,t}_u(1,0) = -\frac{\psi_u + \lambda_u{}^{\mathbf{P}}Y_u}{\mathcal{N}_u(Y)}\,{}^t\mathcal{E}\Big(-\frac{\psi + \lambda\,{}^{\mathbf{P}}Y}{\mathcal{N}(Y)}\boldsymbol{\cdot} S\Big)_{u-}, \qquad t \le u \le T, \tag{2.19}$$

and $q = V^0(1)$ is the unique bounded strictly positive solution of (2.18).

2) *Suppose in addition that there is some $Q \in I\!P^2_{\mathrm{e},\sigma}(S)$ satisfying the reverse Hölder inequality $R_2(P)$. Then $q = V^0(1)$ is the unique solution to the BSDE (2.18) in the class of processes satisfying $c \le Y \le C$ for positive constants c, C. Moreover, the optimal $\vartheta^{*,t}(1,0)$ exist and are given by (2.19).*

Proof. Throughout this proof, we write $\vartheta^{*,t}$ for $\vartheta^{*,t}(1,0)$ and denote by m a generic local P-martingale that can change from one appearance to the next.

1) For part 1), we start by deriving the BSDE (2.18). By part 1) of Lemma 1.5, $q = v^{(2)}$ is a P-submartingale, hence a P-special semimartingale with canonical decomposition $q = q_0 + N^q + B^q$, and $0 \le q \le 1$ implies that $q \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and N^q has bounded jumps and is in $\mathcal{M}^2_{0,\mathrm{loc}}(P)$. The Galtchouk–Kunita–Watanabe decomposition thus allows us to write

$$q = q_0 + \varphi\boldsymbol{\cdot} M + L^q + B^q \tag{2.20}$$

with $\varphi \in L^2_{\mathrm{loc}}(M)$ and $L^q \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to M. Combining this with (2.7) and Yoeurp's lemma then gives

$$[q, S] = m + \varphi\boldsymbol{\cdot}[M] + [A, B^q] = m + (\varphi + \lambda\Delta B^q)\boldsymbol{\cdot}\langle M\rangle. \tag{2.21}$$

We now apply Itô's formula to the process $X^\vartheta_{t,u} := x + \int_t^u \vartheta_r\,dS_r$, $t \le u \le T$, for $x \in I\!R$, $t \in [0,T]$ and $\vartheta \in \Theta$. (We sometimes omit writing the dependence of X^ϑ on t.) This gives

$$(X^\vartheta_u)^2 = x^2 + 2\int_t^u X^\vartheta_{r-}\vartheta_r\,dS_r + \int_t^u \vartheta^2_r\,d[S]_r. \tag{2.22}$$

Next we apply the product rule with (2.22), (2.20), (2.7), (2.21) and then use $A = \int \lambda\,d\langle M\rangle$ and $q_-\boldsymbol{\cdot}[S] + [q,[S]\,] = (q_- + \Delta q)\boldsymbol{\cdot}[S] = q\boldsymbol{\cdot}[S]$ as well as (2.8), (2.10) for q and (2.15) to obtain

$$\begin{aligned}(X^\vartheta_{t,u})^2 q_u - x^2 q_t &= m_u - m_t + \int_t^u (X^\vartheta_{r-})^2\,dB^q_r + 2\int_t^u q_{r-}X^\vartheta_{r-}\vartheta_r\,dA_r + \int_t^u q_{r-}\vartheta^2_r\,d[S]_r\\ &\quad + 2\int_t^u X^\vartheta_{r-}\vartheta_r(\varphi_r + \lambda_r\Delta B^q_r)\,d\langle M\rangle_r + \int_t^u \vartheta^2_r\,d\big[q,[S]\big]_r\\ &= m_u - m_t + \int_t^u (X^\vartheta_{r-})^2\,dB^q_r\\ &\quad + \int_t^u \big(2X^\vartheta_{r-}\vartheta_r(\varphi_r + \lambda_r{}^{\mathbf{P}}q_r) + \vartheta^2_r\mathcal{N}_r(q)\big)\,d\langle M\rangle_r\\ &= m_u - m_t + \int_t^u f(r, X^\vartheta_{t,r-};\vartheta)\,dC_r,\end{aligned} \tag{2.23}$$

where $C \in \mathcal{A}^+_{\mathrm{loc}}(P)$ is a predictable process with $B^q = \int \beta \, dC$, $\langle M \rangle = \int \nu \, dC$ and

$$f(r, y; \vartheta) := y^2 \beta_r + G_r(y, \vartheta_r)\nu_r := y^2\beta_r + \big(2y\vartheta_r(\varphi_r + \lambda_r {}^{\mathbf{P}}q_r) + \vartheta_r^2 \mathcal{N}_r(q)\big)\nu_r \tag{2.24}$$

is a quadratic polynomial in y with random processes as coefficients. Replacing C_t by $C_t + t$, we can assume that C as well as its continuous part C^c is strictly increasing.

By Corollary 1.2, $((X^\vartheta_{t,u})^2 q_u)_{t \le u \le T}$ is a P-submartingale for every $\vartheta \in \Theta$ and a P-martingale for the optimal $\vartheta^{*,t} \in \Theta$, if that exists. This means that the dC-integral in (2.23) is increasing for every $\vartheta \in \Theta$ and identically 0 for $\vartheta = \vartheta^{*,t}$, and the same then applies separately for the corresponding integrals with respect to the continuous and purely discontinuous parts C^c and C^d of C. Similarly as in Mania/Tevzadze (2003a), we therefore obtain for each $x \in I\!R$

$$\operatorname*{ess\,inf}_{\vartheta \in \Theta} f(r, x; \vartheta) = x^2 \beta_r + \nu_r \operatorname*{ess\,inf}_{\vartheta \in \Theta} G_r(x, \vartheta_r) = 0 \qquad P \otimes C\text{-a.e.}; \tag{2.25}$$

the details for this step are a bit more technical and postponed to step 2). Using the definition of $G_r(y, \vartheta_r)$ in (2.24) and completing the square gives

$$G_r(x, \vartheta_r) = \mathcal{N}_r(q)\Big(\vartheta_r + x \frac{\varphi_r + \lambda_r {}^{\mathbf{P}}q_r}{\mathcal{N}_r(q)}\Big)^2 - x^2 \frac{(\varphi_r + \lambda_r {}^{\mathbf{P}}q_r)^2}{\mathcal{N}_r(q)}, \tag{2.26}$$

and we claim that for a localising sequence $(\tau_n)_{n \in I\!N}$,

$$\vartheta^n := -x \frac{\varphi + \lambda \, {}^{\mathbf{P}}q}{\mathcal{N}(q)} I_{[\![0, \tau_n]\!]} \in \Theta. \tag{2.27}$$

Indeed, $\mathcal{N}(q)$ is locally bounded away from 0 by Lemma 2.3, and ${}^{\mathbf{P}}q$ is bounded like q due to Lemma 1.5. Moreover, $\int \lambda^2 \, d\langle M \rangle$ is locally bounded since it is predictable and RCLL, and φ is locally in $L^2(M)$ by construction. Thus we obtain via Cauchy–Schwarz that both φ and λ, and then also the ratio in (2.27), are locally in $L^2(M) \cap L^2(A) = \Theta$, as claimed. Inserting ϑ^n into (2.26) makes the first term in (2.26) vanish for $n \to \infty$ and thus yields

$$\operatorname*{ess\,inf}_{\vartheta \in \Theta} G_r(x, \vartheta_r) = -x^2 \frac{(\varphi_r + \lambda_r {}^{\mathbf{P}}q_r)^2}{\mathcal{N}_r(q)} \qquad P \otimes C\text{-a.e.}$$

Plugging this into (2.25) and integrating gives $B^q = \int \beta \, dC = \int \frac{(\varphi + \lambda {}^{\mathbf{P}}q)^2}{\mathcal{N}(q)} \, d\langle M \rangle$, and plugging that in turn into (2.20) shows that the triple (q, φ, L^q) solves the BSDE (2.18). Moreover, we see from Lemma 2.1 and $q \le 1$ that q is strictly positive and bounded.

2) To prove (2.25), we use the same basic approach as in Mania/Tevzadze (2003a), but we must be more careful and handle jumps since S is not continuous. For ease of notation, we sometimes omit the third argument ϑ of f. We first write $C = C^c + C^d$ and denote by $(\tau_k)_{k \in I\!N}$ a sequence of stopping times exhausting the jumps of C^d (or C). Each τ_k is

predictable because C is predictable. By Corollary 1.2, we then have with probability 1 that $C_{\cdot}(\omega)$ is RCLL and simultaneously for all rational $s \in [0,T]$ that

$$\int_s^u f(r, X^\vartheta_{s,r-}; \vartheta)\, dC_r,\ s \le u \le T, \quad \text{is increasing,} \tag{2.28}$$

$$\int_s^u f(r, X^\vartheta_{s,r-}; \vartheta)\, dC^c_r,\ s \le u \le T, \quad \text{is increasing,} \tag{2.29}$$

for each $\vartheta \in \Theta$, and for the optimal $\vartheta^{*,s}$, the processes in (2.28) and (2.29) vanish identically. Indeed, (2.29) follows from (2.28) since the process in (2.29) is simply the continuous part of the process in (2.28). For any τ_k, we thus have with probability 1 that

$$\int_s^{\tau_k(\omega)} f(r, X^\vartheta_{s,r-}; \vartheta)(\omega)\, dC_r(\omega) \ge 0 \qquad \text{for all rational } s < \tau_k(\omega).$$

Because τ_k is predictable, there are stopping times $(\sigma^{(n)}_k)_{n\in\mathbb{N}}$ taking only rational values and such that $\lim\limits_{n\to\infty} \sigma^{(n)}_k = \tau_k$ and $\sigma^{(n)}_k < \tau_k$ on $\{\tau_k > 0\} = \Omega$; see Theorem IV.77 in Dellacherie/ Meyer (1978). Thus we obtain for P-almost all ω that

$$\int_{\sigma^{(n)}_k(\omega)}^{\tau_k(\omega)} f\big(r, X^\vartheta_{\sigma^{(n)}_k, r-}; \vartheta\big)(\omega)\, dC_r(\omega) \ge 0 \qquad \text{for all } k \text{ and } n.$$

As $n \to \infty$, these integrals tend to $f(\tau_k, X^\vartheta_{\tau_k-,\tau_k-}; \vartheta)(\omega)\Delta C_{\tau_k}(\omega) = f(\tau_k, x; \vartheta)(\omega)\Delta C_{\tau_k}(\omega)$ because $X^\vartheta_{\tau_k-,\tau_k-} = x$, and so we get

$$f(\tau_k, x; \vartheta)\Delta C_{\tau_k} \ge 0 \qquad \text{for all } k \in \mathbb{N},\ P\text{-a.s.,} \tag{2.30}$$

which means that $f(\,\cdot\,, x; \vartheta) \ge 0$ $P \otimes C^d$-a.e., for each $\vartheta \in \Theta$. For the optimal $\vartheta^{*,s}$, we get the null process in (2.28), hence equality in (2.30), and so we have

$$\operatorname*{ess\,inf}_{\vartheta\in\Theta} f(\,\cdot\,, x; \vartheta) = 0 \qquad P \otimes C^d\text{-a.e.} \tag{2.31}$$

For the continuous part C^c, (2.29) gives with $\tau_s(\varepsilon) := \inf\{t \ge s \,|\, C^c_t \ge C^c_s + \varepsilon\}$ that

$$\int_s^{\tau_s(\varepsilon)} f(t, X^\vartheta_{s,t-}; \vartheta)\, dC^c_t \ge 0 \quad \text{for all rational } s \in [0,T],\ P\text{-a.s.} \tag{2.32}$$

We claim that for each $u \ge s$,

$$s \mapsto \int_s^u f(t, X^\vartheta_{s,t-}; \vartheta)\, dC^c_t \quad \text{is } P\text{-a.s. right-continuous.} \tag{2.33}$$

Postponing the argument for the moment, we obtain that the inequality in (2.32) also holds for all $s \in [0,T]$, P-a.s. Setting $\sigma_t(\varepsilon) := \inf\{s \geq 0 \,|\, C^c_s \geq C^c_t - \varepsilon\}$, we then get as in Appendix B of Mania/Tevzadze (2003a) via Fubini's theorem that (dropping arguments ϑ from f)

$$
\begin{aligned}
(2.34)\qquad \int_0^T \left| \frac{1}{\varepsilon} \int_s^{\tau_s(\varepsilon)} f(t, X^\vartheta_{s,t-})\, dC^c_t - f(s,x) \right| dC^c_s &\leq \int_0^T \frac{1}{\varepsilon} \int_{\sigma_t(\varepsilon)}^t |f(t, X^\vartheta_{s,t-}) - f(t,x)|\, dC^c_s\, dC^c_t \\
&\quad + \int_0^T \frac{1}{\varepsilon} \int_s^{\tau_s(\varepsilon)} |f(t,x) - f(s,x)|\, dC^c_t\, dC^c_s;
\end{aligned}
$$

this uses that $C^c_{\tau_s(\varepsilon)} - C^c_s = \varepsilon$ by continuity of C^c. The second term on the right-hand side of (2.34) tends to 0 as $\varepsilon \searrow 0$ by Corollary B.1 in Mania/Tevzadze (2003a). Writing

$$
b^\varepsilon_t := \sup\left\{ |X^\vartheta_{s,t-} - x| \,\middle|\, \sigma_t(\varepsilon) < s < t \right\} = \sup\left\{ \left| \int_s^{t-} \vartheta_r\, dS_r \right| \,\middle|\, \sigma_t(\varepsilon) < s < t \right\},
$$

we have $\sigma_t(\varepsilon) \nearrow t$ for $\varepsilon \searrow 0$ by continuity of C^c and therefore $b^\varepsilon_t \searrow 0$ as $\varepsilon \searrow 0$. Moreover, we have (uniformly in ε and t) $b^\varepsilon_t \leq 2 \sup\limits_{0 \leq r \leq T} |\vartheta \cdot S_r|$ which is in $L^2(P)$, hence P-a.s. finite, for $\vartheta \in \Theta$. The first term on the right-hand side of (2.34) can now be estimated above by

$$
\int_0^T \sup\left\{ |f(t,y;\vartheta) - f(t,x;\vartheta)| \,\middle|\, |y-x| \leq b^\varepsilon_t \right\} dC^c_t =: \int_0^T h_\varepsilon(t;\vartheta)\, dC^c_t
$$

since $C^c_t - C^c_{\sigma_t(\varepsilon)} = \varepsilon$ by continuity of C^c. Now we use the definition of f in (2.24) to obtain

$$
h_\varepsilon(t;\vartheta) \leq (b^\varepsilon_t)^2 |\beta_t| + b^\varepsilon_t \big(2|\beta_t||x| + 2\nu_t|\vartheta_t|(|\varphi_t| + |\lambda_t|\, {}^{\mathrm{P}}q_t) \big).
$$

This shows that P-a.s., $h_\varepsilon(t;\vartheta) \to 0$ for all t as $\varepsilon \searrow 0$. Moreover, b^ε_t can be bounded uniformly in ε and t, P-a.s., and using

$$
(2.35)\qquad \int_0^T |\beta_t|\, dC^c_t \leq \int_0^T |dB^q_t|,
$$

$$
(2.36)\qquad \int_0^T \nu_t |\vartheta_t| (|\varphi_t| + |\lambda_t|\, {}^{\mathrm{P}}q_t)\, dC^c_t \leq \left(\int_0^T \vartheta_t^2\, d\langle M\rangle_t \right)^{\frac{1}{2}} \left(2 \int_0^T (\varphi_t^2 + \lambda_t^2)\, d\langle M\rangle_t \right)^{\frac{1}{2}}
$$

shows that we can apply dominated convergence to get $\int_0^T h_\varepsilon(t;\vartheta)\, dC^c_t \longrightarrow 0$ as $\varepsilon \searrow 0$, P-a.s. With a similar argument, we can prove (2.33). Indeed, for $s_n \searrow s$, we have

$$
\begin{aligned}
\left| \int_s^u f(t, X^\vartheta_{s,t-})\, dC^c_t - \int_{s_n}^u f(t, X^\vartheta_{s_n,t-})\, dC^c_t \right| &\leq \int_s^{s_n} |f(t, X^\vartheta_{s,t-})|\, dC^c_t \\
&\quad + \int_{s_n}^u |f(t, X^\vartheta_{s,t-}) - f(t, X^\vartheta_{s_n,t-})|\, dC^c_t
\end{aligned}
$$

and the first term on the right-hand side tends to 0 P-a.s. as $n \to \infty$ by continuity of C^c. Writing $h_n(t) := |f(t, X^\vartheta_{s,t-}) - f(t, X^\vartheta_{s_n,t-})|$, we have $h_n(t) \to 0$ as $n \to \infty$ by the right-continuity of the stochastic integral and since f from (2.24) is continuous with respect to the second argument y. So (2.33) will follow by dominated convergence as soon as we show that

$$\int_0^T \sup_{n \in \mathbb{N}} h_n(t)\, dC^c_t < \infty \qquad P\text{-a.s.} \tag{2.37}$$

But the definition of f in (2.24) yields that

$$h_n(t) \le 4|\beta_t|\big(|x|^2 + \sup_{0 \le r \le T} |\vartheta \bullet S_r|^2\big) + 2|\nu_t||\vartheta_t|(|\varphi_t| + |\lambda_t|^{\mathbf{P}} q_t) \sup_{0 \le r \le T} |\vartheta \bullet S_r|,$$

and so (2.37) follows again by (2.35) and (2.36) because $\sup_{0 \le r \le T} |\vartheta \bullet S_r| < \infty$ P-a.s. This establishes (2.33).

Putting together all the results so far, (2.34) therefore yields that with probability 1, we have $\frac{1}{\varepsilon} \int_s^{\tau_s(\varepsilon;\vartheta)} f(t, X^\vartheta_{s,t-}; \vartheta)\, dC^c_t \longrightarrow f(s,x;\vartheta)$ in $L^1(dC^c)$ as $\varepsilon \searrow 0$. Together with (2.32), this gives $f(\,\cdot\,, x; \vartheta) \ge 0$ $P \otimes C^c$-a.e., for each $\vartheta \in \Theta$. For the optimal $\vartheta^{*,s}$, we again get equality so that finally

$$\operatorname*{ess\,inf}_{\vartheta \in \Theta} f(\,\cdot\,, x; \vartheta) = 0 \qquad P \otimes C^c\text{-a.e.},$$

and combining this with (2.31) yields (2.25).

3) We next show that $\vartheta^{*,t}$ for fixed t is given by (2.19). Since (q, φ, L^q) satisfies (2.18), Itô's formula gives via (2.22) and (2.8)–(2.11) like in (2.23) for any $\vartheta \in \Theta$ that

$$\begin{aligned}(X^\vartheta_u)^2 q_u - x^2 q_t = m_u - m_t + \int_t^u \Bigg(&(X^\vartheta_{r-})^2 \frac{(\varphi_r + \lambda_r{}^{\mathbf{P}} q_r)^2}{\mathcal{N}_r(q)} + 2q_{r-} X^\vartheta_{r-} \vartheta_r \lambda_r \\ &+ q_{r-}\vartheta_r^2 (1 + \lambda_r^2 \Delta\langle M\rangle_r) + 2X^\vartheta_{r-}\vartheta_r(\varphi_r + \lambda_r \Delta B^q_r) \\ &+ \vartheta_r^2\big(\Delta B^q_r(1 + \lambda_r^2 \Delta\langle M\rangle_r) + g_r(q)\big)\Bigg)\, d\langle M\rangle_r \\ = m_u - m_t + \int_t^u &\left(\vartheta_r \sqrt{\mathcal{N}_r(q)} + X^\vartheta_{r-} \frac{\varphi_r + \lambda_r{}^{\mathbf{P}} q_r}{\sqrt{\mathcal{N}_r(q)}} \right)^2 d\langle M\rangle_r.\end{aligned} \tag{2.38}$$

By Corollary 1.2, the process in (2.38) is a martingale on $[\![t, T]\!]$ for the optimal $\vartheta^{*,t}$, and so

$$\vartheta^{*,t} = -X^{\vartheta^{*,t}}_{-} \frac{\varphi + \lambda\, {}^{\mathbf{P}} q}{\mathcal{N}(q)} \qquad P \otimes \langle M\rangle\text{-a.e. on }]\!]t, T]\!]. \tag{2.39}$$

Integrating with respect to S thus shows for $x = 1$ that $X^{\vartheta^{*,t}} = 1 + \int_t^{\cdot} \vartheta^{*,t}\,dS$ satisfies the linear SDE $X_u^{\vartheta^{*,t}} = 1 - \int_t^u X_{r-}^{\vartheta^{*,t}} \frac{\varphi_r + \lambda_r^{\mathbf{P}} q_r}{\mathcal{N}_r(q)}\,dS_r$ for $t \le u \le T$, and this implies that

$$X^{\vartheta^{*,t}} = {}^t\mathcal{E}\Big(-\frac{\varphi + \lambda^{\mathbf{P}} q}{\mathcal{N}(q)} \centerdot S \Big). \tag{2.40}$$

Because $\vartheta^{*,t}$ is in Θ, we have $X^{\vartheta^{*,t}} \in \mathcal{S}^2(P)$ so that the stochastic exponential is indeed in $\mathcal{S}^2(P)$; and plugging (2.40) into (2.39) yields the expression (2.19) for $\vartheta^{*,t}$. Since t was arbitrary, we have now shown that a) implies b) and that we then have (2.19).

4) Conversely, let us start from b). Again fix t. Using the fact that (Y, ψ, L) solves the BSDE (2.18), we obtain completely analogously as for (2.38) for any $\vartheta \in L(S)$ that

$$(X_u^{\vartheta})^2 Y_u - x^2 Y_t = m_u - m_t + \int_t^u \left(\vartheta_r \sqrt{\mathcal{N}_r(Y)} + X_{r-}^{\vartheta} \frac{\psi_r + \lambda_r^{\mathbf{P}} Y_r}{\sqrt{\mathcal{N}_r(Y)}} \right)^2 d\langle M \rangle_r \tag{2.41}$$

for $t \le u \le T$. So $(X^{\vartheta})^2 Y$ is a local P-submartingale on $[\![t, T]\!]$; but since Y is bounded and $1 + \vartheta \centerdot S \in \mathcal{S}^2(P)$ for $\vartheta \in \Theta$, we get that $(X^{\vartheta})^2 Y$ is actually a true P-submartingale on $[\![t, T]\!]$ so that $Y_T = 1$ gives $Y_t \le E\Big[\Big(1 + \int_t^T \vartheta_r\,dS_r\Big)^2 \Big| \mathcal{F}_t\Big]$ for any $\vartheta \in \Theta$. The definition in (2.1) thus yields $Y_t \le V_t^0(1) = q_t$ for all $t \in [0, T]$. To prove the converse inequality, define the predictable process $\widetilde{\vartheta}^{(t)}$ by the right-hand side of (2.19). Integrating then shows as for (2.40) that $X^{\widetilde{\vartheta}^{(t)}} = {}^t\mathcal{E}(-\frac{\psi + \lambda^{\mathbf{P}} Y}{\mathcal{N}(Y)} \centerdot S)$, and because this stochastic exponential is in $\mathcal{S}^2(P)$ by the assumption in b), we see that $\widetilde{\vartheta}^{(t)}$ coming from (2.19) is actually in Θ. Plugging $\widetilde{\vartheta}^{(t)}$ into (2.41) shows by (2.19) that the $d\langle M \rangle$-integral vanishes; so $(X^{\widetilde{\vartheta}^{(t)}})^2 Y$ is a P-martingale on $[\![t, T]\!]$ and hence $Y_t = E\Big[\Big(1 + \int_t^T \widetilde{\vartheta}_r^{(t)}\,dS_r\Big)^2 \Big| \mathcal{F}_t\Big] \ge V_t^0(1) = q_t$ by (2.1). So we obtain $Y = q$, hence also $\psi \centerdot M = \varphi \centerdot M$, $L = L^q$, and this shows that any solution of (2.18) with the properties in b) coincides with (q, φ, L^q), giving uniqueness. Finally, $Y = q$ shows that $(X^{\vartheta})^2 q$ is a P-submartingale on $[\![t, T]\!]$ for any $\vartheta \in \Theta$ and a P-martingale for $\vartheta = \widetilde{\vartheta}^{(t)} \in \Theta$; so $\widetilde{\vartheta}^{(t)}$ is optimal by Corollary 1.2 and in particular, an optimal $\vartheta^{*,t}(1, 0) = \widetilde{\vartheta}^{(t)}$ exists. Since t was arbitrary, we have also shown that b) implies a), and part 1) of Theorem 2.4 is proved.

5) It remains to prove part 2). But if there is some $Q \in \mathbb{P}^2_{e,\sigma}(S)$ with $R_2(P)$, the space $L^2(\mathcal{F}_t, P) + G_{t,T}(\Theta) = \{X + \vartheta \centerdot S_T \mid X \in L^2(\mathcal{F}_t, P), \vartheta \in \Theta_{t,T}\}$ is closed in $L^2(P)$ by Theorem 5.2 of Choulli/Krawczyk/Stricker (1998), for every t, so that an optimal $\vartheta^{*,t}$ exists. Moreover, we then have $q \ge \delta > 0$ by Lemma 2.1, and so part 2) follows directly from part 1). **q.e.d.**

Remark. If $d > 1$, the backward equation (2.18) looks more complicated. Using the notation from the remark before Theorem 2.4, in particular (2.16) and (2.17), the equation reads

$$Y_t = Y_0 + \int_0^t (\psi_s + \lambda_s {}^{\mathbf{P}}Y_s)^{\mathrm{tr}} \mu_s \big(\mathcal{N}_s(Y)\big)^{-1} \mu_s(\psi_s + \lambda_s {}^{\mathbf{P}}Y_s)\, dB_s + \int_0^t \psi_s\, dM_s + L_t, \qquad Y_T = 1,$$

where $\mathcal{N}_s(Y) := {}^{\mathbf{P}}Y_s\,(\mu_s + (\mu_s\lambda_s)^{\mathrm{tr}}\mu_s\lambda_s\Delta B_s) + g_s(Y)$. We do not give details. ◇

For later use, we record the following consequence of Theorem 2.4.

Corollary 2.5. *Under the assumptions of Theorem 2.4, suppose a) or b) there hold. Define*

$$\gamma := -\frac{\psi + \lambda\, {}^{\mathbf{P}}Y}{\mathcal{N}(Y)} = -\frac{\psi + \lambda\, {}^{\mathbf{P}}Y}{{}^{\mathbf{P}}Y(1 + \lambda^2 \Delta\langle M\rangle) + g(Y)}, \tag{2.42}$$

where (Y, ψ, L) is the solution of the BSDE (2.18), and recall the process $v^{(1)}$ from the quadratic representation (1.6) of V^H. For every $t \in [0,T]$, we then have

$$v_t^{(1)} = E\big[H\, {}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_T \,\big|\, \mathcal{F}_t\big] \qquad P\text{-a.s.} \tag{2.43}$$

and the process $({}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_u\, v_u^{(1)})_{t \le u \le T}$ is a P-martingale on $[\![t,T]\!]$.

Proof. Fix t. Because we have $1 + \int_t^T \vartheta_r^{*,t}(1,0)\, dS_r = X_T^{\vartheta^{*,t}} = {}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_T$ by (2.40) and the definition (2.42) of γ, (2.43) follows directly from (1.8). Moreover, it is easy to check that for any semimartingale X and any $u \le T$, we have ${}^u\mathcal{E}(X)_T = \frac{\mathcal{E}(X)_T}{\mathcal{E}(X)_u}$ P-a.s. on $\{\mathcal{E}(X)_u \neq 0\}$ and $\mathcal{E}(X)_T = 0$ P-a.s. on $\{\mathcal{E}(X)_u = 0\}$. Taking $X := \gamma \boldsymbol{\cdot} S - (\gamma \boldsymbol{\cdot} S)^t$, $u \ge t$ and setting for brevity $D_u := \{{}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_u \neq 0\}$ therefore gives the desired martingale property via

$$\begin{aligned} {}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_u\, v_u^{(1)} &= I_{D_u}\, {}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_u\, E\big[H\, {}^u\mathcal{E}(\gamma \boldsymbol{\cdot} S)_T \,\big|\, \mathcal{F}_u\big] \\ &= I_{D_u} E\big[H\, {}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_T \,\big|\, \mathcal{F}_u\big] \\ &= E\big[H\, {}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S)_T \,\big|\, \mathcal{F}_u\big]; \end{aligned}$$

integrability holds since $H \in L^2(P)$ and ${}^t\mathcal{E}(\gamma \boldsymbol{\cdot} S) \in \mathcal{S}^2(P)$ by part 1b) of Theorem 2.4. **q.e.d.**

As before, we can connect our results to the dual problem, as follows.

Proposition 2.6. *Under the assumptions of Theorem 2.4, suppose a) or b) there hold. Then the variance-optimal signed martingale measure $\widetilde{Q} \in \mathbb{P}^2_{\mathrm{s},\sigma}(S)$ is given by*

$$\frac{d\widetilde{Q}}{dP} = \frac{1}{Y_0}\mathcal{E}\Big(-\frac{\psi + \lambda\, {}^{\mathbf{P}}Y}{\mathcal{N}(Y)} \boldsymbol{\cdot} S\Big)_T = \frac{1}{Y_0}\mathcal{E}(\gamma \boldsymbol{\cdot} S)_T, \tag{2.44}$$

where (Y, ψ, L) is the solution of the BSDE (2.18). If we have in addition that

(2.45) $$\gamma_t \Delta S_t = -\frac{\psi_t + \lambda_t {}^{\mathbf{p}}Y_t}{{}^{\mathbf{p}}Y_t(1 + \lambda_t^2 \Delta\langle M\rangle_t) + g_t(Y)} \Delta S_t > -1 \qquad P\text{-a.s. for } 0 \leq t \leq T,$$

then the VOMM exists and is given by $\widetilde{Q}$ from (2.44).

Proof. From the BSDE (2.18) and Itô's formula, we obtain by straightforward computation that the product $Y\mathcal{E}(-\frac{\psi+\lambda {}^{\mathbf{p}}Y}{\mathcal{N}(Y)} \cdot S)$ is a local P-martingale. But it is even a true P-martingale since Y is bounded and the stochastic exponential is in $\mathcal{S}^2(P)$, and so (2.44) defines a signed measure $\widetilde{Q} \ll P$ with P-square-integrable density process $Z^{\widetilde{Q}} = \frac{Y}{Y_0}\mathcal{E}(-\frac{\psi+\lambda {}^{\mathbf{p}}Y}{\mathcal{N}(Y)} \cdot S)$ and $\widetilde{Q}[\Omega] = 1$. Note for (2.44) that $Y_T = 1$. Another straightforward but slightly lengthier computation shows that $Z^{\widetilde{Q}}S$ is a local P-martingale so that $\widetilde{Q} \in I\!\!P^2_{\mathrm{s},\sigma}(S)$. Finally, the representation (2.44) of $\frac{d\widetilde{Q}}{dP}$ as a constant plus a "good" stochastic integral of S implies that $\widetilde{Q}$ is variance-optimal; see for instance Lemma 2.1 in Delbaen/Schachermayer (1996). Note here that the same argument as in step 4) of the proof of Theorem 2.4 implies that the integrand $\vartheta := \frac{1}{Y_0}\gamma\mathcal{E}(\gamma \cdot S)_-$ is in Θ so that $\vartheta \cdot S$ is a Q-martingale for every $Q \in I\!\!P^2_{\mathrm{e},\sigma}(S)$. If (2.45) holds, then clearly $Z^{\widetilde{Q}} > 0$; so $\widetilde{Q}$ is then equivalent to P, hence in $I\!\!P^2_{\mathrm{e},\sigma}(S)$, and is the VOMM. **q.e.d.**

Remark. From (2.43), the proof of Proposition 2.6 and $Y = v^{(2)}$, we can see that under the assumptions of Theorem 2.4 and (2.45), the process $v^{(1)}\mathcal{E}(\gamma \cdot S) = v^{(1)}Y_0 Z^{\widetilde{Q}}/Y$ is a P-martingale with final value $H\mathcal{E}(\gamma \cdot S)_T = HY_0 Z^{\widetilde{Q}}_T$. This implies that

$$\frac{v_t^{(1)}}{v_t^{(2)}} = \frac{v_t^{(1)}}{Y_t} = E_{\widetilde{Q}}[H \,|\, \mathcal{F}_t], \qquad 0 \leq t \leq T.$$

◇

3. Mean-variance hedging: from $(1, 0)$ to (x, H)

Recall from Theorem 1.4 that the dynamic value process of the mean-variance hedging problem has the quadratic form

$$V^H(x) = v^{(0)} - 2v^{(1)}x + v^{(2)}x^2.$$

Our goals in this section are to describe the coefficient processes $v^{(0)}, v^{(1)}, v^{(2)}$ via backward stochastic differential equations (BSDEs) and to give explicit expressions for the optimal strategies $\vartheta^{*,t}(x, H)$. This will be done under the same assumptions as in Section 2.

A general solution for the MVH problem has been given by Černý/Kallsen (2007) in their Theorem 4.10 and Corollary 4.11. However, that solution involves either a process N which is very hard to find (see Černý/Kallsen (2007), Definition 3.12) or the variance-optimal martingale measure (called Q^* in Černý/Kallsen (2007), see their Proposition 3.13) which is also notoriously difficult to determine. With our approach, we can be more explicit.

To formulate our main result, we introduce the system of BSDEs

$$(3.1)\ dY_s^{(2)} = \frac{(\psi_s^{(2)} + \lambda_s {}^{\mathbf{P}}Y_s^{(2)})^2}{\mathcal{N}_s(Y^{(2)})}\, d\langle M\rangle_s + \psi_s^{(2)}\, dM_s + dL_s^{(2)}, \qquad Y_T^{(2)} = 1,$$

$$(3.2)\ dY_s^{(1)} = \frac{(\psi_s^{(2)} + \lambda_s {}^{\mathbf{P}}Y_s^{(2)})(\psi_s^{(1)} + \lambda_s {}^{\mathbf{P}}Y_s^{(1)})}{\mathcal{N}_s(Y^{(2)})}\, d\langle M\rangle_s + \psi_s^{(1)}\, dM_s + dL_s^{(1)}, \qquad Y_T^{(1)} = H,$$

$$(3.3)\ dY_s^{(0)} = \frac{(\psi_s^{(1)} + \lambda_s {}^{\mathbf{P}}Y_s^{(1)})^2}{\mathcal{N}_s(Y^{(2)})}\, d\langle M\rangle_s + dN_s^{(0)}, \qquad Y_T^{(0)} = H^2.$$

A *solution* of this system consists of tuples $(Y^{(2)}, \psi^{(2)}, L^{(2)})$, $(Y^{(1)}, \psi^{(1)}, L^{(1)})$, $(Y^{(0)}, N^{(0)})$ where $\psi^{(2)}, \psi^{(1)}$ are in $L^1_{\mathrm{loc}}(M)$; $L^{(2)}, L^{(1)}$ are in $\mathcal{M}_{0,\mathrm{loc}}(P)$ and strongly P-orthogonal to M; $N^{(0)}$ is a local P-martingale; and $Y^{(2)}, Y^{(1)}, Y^{(0)}$ are P-special semimartingales with $[N^{Y^{(2)}}, [S]\,] \in \mathcal{A}_{\mathrm{loc}}(P)$. We point out that (3.1) is the same equation as (2.18) before Theorem 2.4. Note also that (given $Y^{(2)}, \psi^{(2)}, L^{(2)}$) the equation (3.2) is linear and can therefore be solved explicitly; and $Y^{(0)}$ and $N^{(0)}$ for (3.3) can even be written down directly. In the case where S is continuous, this system has been obtained and studied in Mania/Tevzadze (2003a) or (under the additional assumption that $I\!F$ is continuous) in Bobrovnytska/Schweizer (2004). For a Markovian setting within a Brownian filtration, the corresponding PDEs can also be found in Bertsimas/Kogan/Lo (2001), with a heuristic treatment.

Theorem 3.1. *Suppose (as in Theorem 2.4) that $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$, and fix $H \in L^2(\mathcal{F}_T, P)$. Then:*

1) *The following two assertions are equivalent:*

a) *For every $t \in [0,T]$, there exists an optimal $\vartheta^{*,t}(x,H) \in \Theta_{t,T}(0)$ for (1.2) for every $x \in I\!R$.*

b) *For each $x \in I\!R$, there is a solution to the BSDE system (3.1)–(3.3) with*

(i) $L^{(2)} \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$, $\psi^{(2)} \in L^2_{\mathrm{loc}}(M)$, $Y^{(2)}$ bounded and strictly positive, and such that for every $t \in [0,T]$, the process $({}^t\mathcal{E}(-\frac{\psi^{(2)}+\lambda\, {}^{\mathbf{P}}Y^{(2)}}{\mathcal{N}(Y^{(2)})} \cdot S)_u)_{t\le u\le T}$ is in $\mathcal{S}^2(P)$;

(ii) $L^{(1)} \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$, $\psi^{(1)} \in L^2_{\mathrm{loc}}(M)$, $|Y^{(1)}|^2$ of class (D), and such that for every $t \in [0,T]$, the solution $X^{(t)}$ of the linear SDE

$$(3.4)\qquad X_u^{(t)} = x + \int_t^u \frac{\psi_r^{(1)} + \lambda_r {}^{\mathbf{P}}Y_r^{(1)}}{\mathcal{N}_r(Y^{(2)})}\, dS_r - \int_t^u \frac{\psi_r^{(2)} + \lambda_r {}^{\mathbf{P}}Y_r^{(2)}}{\mathcal{N}_r(Y^{(2)})} X_{r-}^{(t)}\, dS_r$$

on $[\![t,T]\!]$ is in $\mathcal{S}^2(P)$;

(iii) $Y^{(0)}$ is a true P-submartingale and (hence) of class (D).

If a) or b) hold, then the value process V^H from (1.2) admits the representation

$$V^H(x) = v^{(0)} - 2v^{(1)}x + v^{(2)}x^2, \tag{3.5}$$

where the processes $v^{(2)}, v^{(1)}, v^{(0)}$ satisfy the BSDE system (3.1)–(3.3), and for every $t \in [0,T]$, the optimal wealth process $X_u^{\vartheta^{*,t}} = x + \int_t^u \vartheta_r^{*,t}(x,H)\, dS_r$, $t \le u \le T$, satisfies the SDE (3.4) and $\vartheta^{*,t} = \vartheta^{*,t}(x,H)$ is given by the feedback formula

$$\vartheta_u^{*,t} = \frac{\psi_u^{(1)} + \lambda_u \mathbf{P} Y_u^{(1)}}{\mathcal{N}_u(Y^{(2)})} - \frac{\psi_u^{(2)} + \lambda_u \mathbf{P} Y_u^{(2)}}{\mathcal{N}_u(Y^{(2)})} X_{u-}^{\vartheta^{*,t}}, \qquad t \le u \le T. \tag{3.6}$$

2) *Suppose in addition that there is some $Q \in I\!P^2_{e,\sigma}(S)$ satisfying the reverse Hölder inequality $R_2(P)$. Then the value process V^H from (1.2) has the form (3.5), where the processes $v^{(2)}, v^{(1)}, v^{(0)}$ are those unique solutions of the BSDE system (3.1)–(3.3) for which $Y^{(0)}$ and $|Y^{(1)}|^2$ are of class (D) and $c \le Y^{(2)} \le C$ for constants $C \ge c > 0$. Moreover, for every $t \in [0,T]$, the optimal strategy $\vartheta^{*,t}(x,H)$ for (1.2) exists, and its wealth process $X^{\vartheta^{*,t}}$ satisfies the SDE (3.4).*

Remark. The integrability condition on the exponential in (i) is not really needed. In fact, like in the proof of Theorem 1.4, one can argue that $\vartheta^{*,t}(1,0) = \vartheta^{*,t}(1,H) - \vartheta^{*,t}(0,H)$ so that the integrability required in (i) follows from that in (ii). But for simpler comparison with Theorem 2.4, we have kept the formulation as a condition. ◇

Proof of Theorem 3.1. As in the proof of Theorem 2.4, we denote by m a generic local P-martingale that can change from one appearance to the next.

1) We first note that as in Theorem 1.4, the existence of optimal strategies $\vartheta^{*,t}(1,0)$ (for $x = 1, H \equiv 0$) follows from a) and is by Theorem 2.4 equivalent to the solvability of (3.1) such that (i) holds in b). So let us start from a), note that (3.5) holds due to Theorem 1.4, and first derive the BSDE for $v^{(1)}$. By Lemma 1.5 and the Galtchouk–Kunita–Watanabe decomposition, we have

$$v^{(1)} = v_0^{(1)} + m^{(1)} + a^{(1)} = v_0^{(1)} + \psi^{(1)} \boldsymbol{\cdot} M + L^{(1)} + a^{(1)} \tag{3.7}$$

with $\psi^{(1)} \in L^2_{\mathrm{loc}}(M)$, $L^{(1)} \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to M, and $a^{(1)}$ predictable and of finite variation. Exactly as for (2.21), this yields

$$[v^{(1)}, S] = m + (\psi^{(1)} + \lambda \Delta a^{(1)}) \boldsymbol{\cdot} \langle M \rangle. \tag{3.8}$$

Now fix t, recall γ from (2.42) in Corollary 2.5 and write $\mathcal{E} := {}^t\mathcal{E}(\gamma\boldsymbol{\cdot} S)$ for brevity. Then combining $d\mathcal{E} = \mathcal{E}_-\gamma\, dS$ with the product rule, (3.7), (2.7), (3.8) and (2.10) yields

$$\begin{aligned}\text{(3.9)}\qquad \mathcal{E}v^{(1)} &= m + \mathcal{E}_-\boldsymbol{\cdot} a^{(1)} + (v^{(1)}_-\mathcal{E}_-\gamma\lambda)\boldsymbol{\cdot}\langle M\rangle + \big(\mathcal{E}_-\gamma(\psi^{(1)} + \lambda\Delta a^{(1)})\big)\boldsymbol{\cdot}\langle M\rangle \\ &= m + \mathcal{E}_-\boldsymbol{\cdot}\Big(a^{(1)} + \big(\gamma(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})\big)\boldsymbol{\cdot}\langle M\rangle\Big).\end{aligned}$$

But we know from Corollary 2.5 that $\mathcal{E}v^{(1)}$ is a P-martingale on $[\![t,T]\!]$, and so the predictable finite variation term on the right-hand side of (3.9) must be identically zero. With $C \in \mathcal{A}^+_{\mathrm{loc}}(P)$ predictable and such that $a^{(1)} \ll C$, $\langle M\rangle \ll C$, we thus obtain that the process $\int {}^t\mathcal{E}(\gamma\boldsymbol{\cdot} S)_-\{\frac{da^{(1)}}{dC} + \gamma(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})\frac{d\langle M\rangle}{dC}\}\, dC$ vanishes identically. Since ${}^t\mathcal{E}(\gamma\boldsymbol{\cdot} S)_t = 1$, we can argue analogously to steps 1) and 2) in the proof of Theorem 2.4 to get

$$\frac{da^{(1)}}{dC} + \gamma(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})\frac{d\langle M\rangle}{dC} = 0 \qquad P\otimes C\text{-a.e.}$$

Integrating with respect to C gives

$$a^{(1)} = -\textstyle\int \gamma(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})\, d\langle M\rangle = \displaystyle\int \frac{(\psi^{(2)} + \lambda\,{}^{\mathbf{P}}Y^{(2)})(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})}{\mathcal{N}(Y^{(2)})}\, d\langle M\rangle,$$

and plugging this into (3.7) shows that $(v^{(1)}, \psi^{(1)}, L^{(1)})$ satisfies the BSDE (3.2). Moreover, as already used, we know from Lemma 1.5 that $|v^{(1)}|^2$ is of class (D), and it only remains for (ii) to check the last integrability property.

2) We next argue that the BSDE (3.3) has a solution, starting with a calculation that is used again later. Fix t, take any ϑ in Θ and consider as in the proof of Theorem 2.4 the process $X^\vartheta_{t,u} := x + \int_t^u \vartheta_r\, dS_r$, $t \le u \le T$. (Again, we usually do not explicitly indicate the dependence of X^ϑ on the starting time t, nor on x.) Lemma 1.5 yields $v^{(0)} = m^{(0)} + a^{(0)}$, and as $v^{(2)}$ satisfies the BSDE (3.1), the same computation as for (2.38) gives with (2.42) that

$$(X^\vartheta_u)^2 v^{(2)}_u - x^2 v^{(2)}_t = m_u - m_t + \int_t^u (\vartheta_r - \gamma_r X^\vartheta_{r-})^2 \mathcal{N}_r(v^{(2)})\, d\langle M\rangle_r.$$

Finally, using the product rule, (2.7), the BSDE (3.2) for $v^{(1)}$, (3.8) and (2.10) leads to

$$\begin{aligned}d(v^{(1)}X^\vartheta) &= v^{(1)}_-\vartheta\, dS + X^\vartheta_-\, dv^{(1)} + \vartheta\, d[v^{(1)}, S] \\ &= dm + v^{(1)}_-\vartheta\lambda\, d\langle M\rangle - X^\vartheta_-\gamma(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})\, d\langle M\rangle + \vartheta(\psi^{(1)} + \lambda\Delta a^{(1)})\, d\langle M\rangle \\ &= dm + (\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})(\vartheta - \gamma X^\vartheta_-)\, d\langle M\rangle.\end{aligned}$$

Using (3.5) and adding up therefore gives

$$\begin{aligned}\text{(3.10)}\qquad V^H_u(X^\vartheta_u) &= v^{(0)}_u - 2v^{(1)}_u X^\vartheta_u + v^{(2)}_u (X^\vartheta_u)^2 \\ &= V^H_t(x) + a^{(0)}_u - a^{(0)}_t - \int_t^u 2(\psi^{(1)}_r + \lambda_r\,{}^{\mathbf{P}}v^{(1)}_r)(\vartheta_r - \gamma_r X^\vartheta_{r-})\, d\langle M\rangle_r \\ &\quad + \int_t^u (\vartheta_r - \gamma_r X^\vartheta_{r-})^2 \mathcal{N}_r(v^{(2)})\, d\langle M\rangle_r + m_u - m_t.\end{aligned}$$

Now choose $x = 0$ and ϑ of the form $\vartheta = yI_{]\!]t,\varrho_t]\!]}$ for some constant $y \in I\!R$, where the stopping time $\varrho_t > t$ is chosen such that ϑ is in Θ; this is possible because S is in $\mathcal{S}^2_{\rm loc}(P)$. Then $\vartheta_r = yI_{\{t<r\le\varrho_t\}}$ and $X^\vartheta_{r-} = y(S_{r-} - S_t)I_{\{t<r\le\varrho_t\}}$, and plugging this into (3.10) and collecting terms gives

$$V^H_u(X^\vartheta_u) - V^H_t(0) = a^{(0)}_u - a^{(0)}_t - 2\int_t^{u\wedge\varrho_t} y(\psi^{(1)}_r + \lambda_r{}^{\mathbf{P}}v^{(1)}_r)\big(1-(S_{r-}-S_t)\gamma_r\big)\,d\langle M\rangle_r$$

$$+\int_t^{u\wedge\varrho_t} y^2\big(1-(S_{r-}-S_t)\gamma_r\big)^2\mathcal{N}_r(v^{(2)})\,d\langle M\rangle_r + m_u - m_t.$$

By Proposition 1.1, this process is always a P-submartingale on $[\![t,T]\!]$. So if we take a predictable $C \in \mathcal{A}^+_{\rm loc}(P)$ with $\langle M\rangle \ll C$ and $a^{(0)} \ll C$, we obtain that the process

$$\int_t^{u\wedge\varrho_t}\left(\Big(y^2\big(1-\gamma_r(S_{r-}-S_t)\big)^2\mathcal{N}_r(v^{(2)}) - 2y(\psi^{(1)}_r + \lambda_r{}^{\mathbf{P}}v^{(1)}_r)\big(1-\gamma_r(S_{r-}-S_t)\big)\Big)\frac{d\langle M\rangle_r}{dC_r} + \frac{da^{(0)}_r}{dC_r}\right)dC_r$$

for $t \le u \le T$ is for all $t \in [0,T]$ and $y \in I\!R$ an increasing process. Again arguing as in steps 1) and 2) of the proof of Theorem 2.4 and using that $S_{r-} - S_s \to 0$ when s increases to r (used for the jumps) or when r decreases to s (used for the continuous part), we get

$$y^2\mathcal{N}(v^{(2)})\frac{d\langle M\rangle}{dC} - 2y(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})\frac{d\langle M\rangle}{dC} + \frac{da^{(0)}}{dC} \ge 0 \qquad \text{for all } y \in I\!R,\ P\otimes C\text{-a.e.}$$

Because $\mathcal{N}(v^{(2)}) > 0$ by Lemma 2.3, we conclude that

$$\text{(3.11)}\qquad \frac{(\psi^{(1)} + \lambda\,{}^{\mathbf{P}}v^{(1)})^2}{\mathcal{N}(v^{(2)})}\frac{d\langle M\rangle}{dC} \le \frac{da^{(0)}}{dC} \qquad P\otimes C\text{-a.e.}$$

This implies that $\int\{da^{(0)} - \frac{(\psi^{(1)}+\lambda\,{}^{\mathbf{P}}v^{(1)})^2}{\mathcal{N}(v^{(2)})}\,d\langle M\rangle\}$ is an increasing process, and since $a^{(0)}$ is P-integrable because $v^{(0)}$ is a P-submartingale by Lemma 1.5, we obtain that

$$E\left[\int_0^T \frac{(\psi^{(1)}_r + \lambda_r{}^{\mathbf{P}}v^{(1)}_r)^2}{\mathcal{N}_r(v^{(2)})}\,d\langle M\rangle_r\right] < \infty.$$

So if we define

$$\text{(3.12)}\qquad Y^{(0)}_t := E\left[H^2 - \int_t^T \frac{(\psi^{(1)}_r + \lambda_r{}^{\mathbf{P}}v^{(1)}_r)^2}{\mathcal{N}_r(v^{(2)})}\,d\langle M\rangle_r\,\middle|\,\mathcal{F}_t\right]$$

$$=: N^{(0)}_t + \int_0^t \frac{(\psi^{(1)}_r + \lambda_r{}^{\mathbf{P}}v^{(1)}_r)^2}{\mathcal{N}_r(v^{(2)})}\,d\langle M\rangle_r,$$

then clearly $(Y^{(0)}, N^{(0)})$ solves (3.3) and $Y^{(0)}$ is a true P-submartingale. This shows that there exists a solution to (3.3) with (iii), but we do not know yet if $v^{(0)} = Y^{(0)}$.

3) To finish the implication "a) $\Longrightarrow$ b)", we now want to prove that each $X^{\vartheta^{*,t}(x,H)}$ satisfies (3.4) and that $v^{(0)} = Y^{(0)}$. We again fix t, take $\vartheta \in \Theta$ and do the same calculation as in (3.10). Completing the square then gives

$$V_u^H(X_u^\vartheta) = V_t^H(x) + m_u - m_t + \int_t^u \left(da_r^{(0)} - \frac{(\psi_r^{(1)} + \lambda_r{}^{\mathbf{P}} v_r^{(1)})^2}{\mathcal{N}_r(v^{(2)})} \, d\langle M \rangle_r \right) + \int_t^u \left((\vartheta_r - \gamma_r X_{r-}^\vartheta)\sqrt{\mathcal{N}_r(v^{(2)})} - \frac{\psi_r^{(1)} + \lambda_r{}^{\mathbf{P}} v_r^{(1)}}{\sqrt{\mathcal{N}_r(v^{(2)})}} \right)^2 d\langle M \rangle_r. \tag{3.13}$$

By Proposition 1.1, this process must be a P-martingale on $[\![t,T]\!]$ if we plug in for ϑ the optimal $\vartheta^{*,t}(x,H)$. Because both integral terms on the right-hand side are increasing due to (3.11), they must then both vanish identically, on $[\![t,T]\!]$ for every t. This firstly gives that

$$a^{(0)} = \int \frac{(\psi^{(1)} + \lambda\, {}^{\mathbf{P}} v^{(1)})^2}{\mathcal{N}(v^{(2)})} \, d\langle M \rangle, \tag{3.14}$$

and as $v^{(0)} = m^{(0)} + a^{(0)}$ is a P-submartingale, comparing (3.12) and (3.14) yields $m^{(0)} = N^{(0)}$, hence $v^{(0)} = Y^{(0)}$, and so $(v^{(0)}, m^{(0)})$ solves the BSDE (3.3) and also is the unique solution satisfying (iii). Secondly, we obtain for the optimal strategy $\vartheta^{*,t} = \vartheta^{*,t}(x,H)$ that

$$\vartheta_u^{*,t} = \frac{\psi_u^{(1)} + \lambda_u{}^{\mathbf{P}} v_u^{(1)}}{\mathcal{N}_u(v^{(2)})} + \gamma_u X_{u-}^{\vartheta^{*,t}},$$

which is (3.6) in view of the definition (2.42) of γ; recall that $(v^{(2)}, \psi^{(2)}, L^{(2)})$ solves (2.18). Integrating with respect to S shows that $X^{\vartheta^{*,t}}$ satisfies the SDE (3.4) on $[\![t,T]\!]$, and since $\vartheta^{*,t}$ is in Θ, the unique solution of (3.4) is in $\mathcal{S}^2(P)$. So we have now proved that a) implies b), and also that we then have (3.5) and (3.6).

4) Conversely, let us start with b); then we have to prove the existence of an optimal $\vartheta^{*,t}(x,H)$. Fix t, set $W_u(x) := Y_u^{(0)} - 2Y_u^{(1)}\, x + Y_u^{(2)}\, x^2$ for $t \le u \le T$ and use (2.22) and the BSDEs (3.1)–(3.3) for $Y^{(2)}, Y^{(1)}, Y^{(0)}$ to compute as for (3.10) and (3.13) that for any $\vartheta \in \Theta$,

$$W_u(X_u^\vartheta) = W_t(x) + m_u - m_t + \int_t^u \left((\vartheta_r - \gamma_r X_{r-}^\vartheta)\sqrt{\mathcal{N}_r(Y^{(2)})} - \frac{\psi_r^{(1)} + \lambda_r{}^{\mathbf{P}} Y_r^{(1)}}{\sqrt{\mathcal{N}_r(Y^{(2)})}} \right)^2 d\langle M \rangle_r \tag{3.15}$$

for $t \le u \le T$. So $W(X^\vartheta)$ is a local P-submartingale on $[\![t,T]\!]$; but we also know from b) that $Y^{(0)}$ is of class (D), $Y^{(2)}$ is bounded and $|Y^{(1)}|^2$ is of class (D). Since X^ϑ is in $\mathcal{S}^2(P)$ for every $\vartheta \in \Theta$, we see that $W(X^\vartheta)$ is thus of class (D), hence a true P-submartingale, and so

$$W_t(x) \le E\big[W_T(X_T^\vartheta) \,\big|\, \mathcal{F}_t\big] = E\Big[\Big(H - x - \int_t^T \vartheta_r \, dS_r\Big)^2 \,\Big|\, \mathcal{F}_t\Big]$$

for any $\vartheta \in \Theta$. This yields $W_t(x) \le V_t^H(x)$ by (1.2). Conversely, if we take the solution $X^{(t)}$ of (3.4) and define

$$\widetilde{\vartheta}^{(t)} := \frac{\psi^{(1)} + \lambda\,{}^{\mathbf{P}}Y^{(1)}}{\mathcal{N}(Y^{(2)})} - \frac{\psi^{(2)} + \lambda\,{}^{\mathbf{P}}Y^{(2)}}{\mathcal{N}(Y^{(2)})} X_-^{(t)},$$

then integrating with respect to S shows that $X^{\widetilde{\vartheta}^{(t)}} = x + \int_t^{\cdot} \widetilde{\vartheta}_r^{(t)}\, dS_r$ equals $X^{(t)}$, since both satisfy (3.4), and is in $\mathcal{S}^2(P)$ due to b) so that $\widetilde{\vartheta}^{(t)}$ is in Θ. Moreover, plugging in $\widetilde{\vartheta}^{(t)}$ for ϑ shows like for (3.15) that $W(X^{\widetilde{\vartheta}^{(t)}})$ is a (true) P-martingale on $[\![t,T]\!]$. This implies that

$$W_t(x) = E\Big[\Big(H - x - \int_t^T \widetilde{\vartheta}_r^{(t)}\, dS_r\Big)^2 \,\Big|\, \mathcal{F}_t\Big] \ge V_t^H(x),$$

and so we conclude that $W_t(x) = V_t^H(x)$ and that $\widetilde{\vartheta}^{(t)}$ is optimal for (1.2), giving existence of $\vartheta^{*,t}(x,H) := \widetilde{\vartheta}^{(t)}$. This proves that b) implies a) and that we then also have $W(x) = V^H(x)$ for all x, hence $Y^{(i)} = v^{(i)}$ for $i = 0,1,2$. This ends the proof of 1).

5) Finally, the assertion of part 2) follows like in Theorem 2.4 from the proof of part 1); we only need to notice again that $L^2(\mathcal{F}_t, P) + G_{t,T}(\Theta)$ is closed in $L^2(P)$ for every t. **q.e.d.**

4. Alternative versions for the BSDEs

In this section, we give equivalent alternative versions for the BSDEs obtained in Sections 2 and 3. One reason is that in some models, these versions are more convenient to work with; a second is that it allows us to discuss how our results relate to existing literature.

For reasons of space, we only look at (2.18) or (3.1) in detail; this is the most complicated equation. **Throughout this section, we assume as in Theorem 2.4 that $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$.** For convenience, we recall that (2.18) reads

$$Y_t = Y_0 + \int_0^t \frac{(\psi_s + \lambda_s {}^{\mathbf{P}}Y_s)^2}{\mathcal{N}_s(Y)}\, d\langle M\rangle_s + \int_0^t \psi_s\, dM_s + L_t, \qquad Y_T = 1, \tag{4.1}$$

where $\mathcal{N}(Y) = {}^{\mathbf{P}}Y(1 + \lambda^2 \Delta\langle M\rangle) + g(Y)$ and $g(Y) = \frac{d[N^Y,[S]]^{\mathbf{P}}}{d\langle M\rangle}$ as in (2.12) and (2.11). A solution of (4.1) is a priori a tuple (Y,ψ,L) with $L \in \mathcal{M}_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to M, $\psi \in L^1_{\mathrm{loc}}(M)$, and Y a P-special semimartingale such that $[N^Y,[S]] \in \mathcal{A}_{\mathrm{loc}}(P)$. In view of Theorem 2.4 (where Y is bounded), we restrict ourselves to solutions with $\psi \in L^2_{\mathrm{loc}}(M)$ and $L \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$. For better comparison with (3.1), we really ought to write a superscript $^{(2)}$ for Y,ψ,L, but we omit this to alleviate the notation.

4.1. Working with M^d

The BSDE (4.1) is written with the local P-martingale M from the canonical decomposition $S = S_0 + M + A = S_0 + M + \int \lambda \, d\langle M\rangle$ of S. In simple models with jumps, it is useful to split $M = M^c + M^d$ into its continuous and purely discontinuous local martingale parts M^c and M^d, respectively. Then $\langle M\rangle = \langle M^c\rangle + \langle M^d\rangle$, and we define the predictable processes

$$\delta^c := \frac{d\langle M^c\rangle}{d\langle M\rangle}, \qquad \delta^d := \frac{d\langle M^d\rangle}{d\langle M\rangle} = 1 - \delta^c.$$

We now consider the backward equation

$$Y_t = Y_0 + \int_0^t \frac{\big(\psi^c_s\delta^c_s + \psi^d_s(1-\delta^c_s) + \lambda_s {}^{\mathbf{P}}Y_s\big)^2}{{}^{\mathbf{P}}Y_s(1+\lambda_s^2\Delta\langle M\rangle_s) + g_s(Y)}\, d\langle M\rangle_s + \int_0^t \psi^c_s \, dM^c_s + \int_0^t \psi^d_s \, dM^d_s + L'_t, \tag{4.2}$$

$$Y_T = 1.$$

A *solution* of (4.2) is a priori a tuple (Y, ψ^c, ψ^d, L') with $L' \in \mathcal{M}_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to both M^c and M^d, $\psi^c \in L^2_{\mathrm{loc}}(M^c)$, $\psi^d \in L^1_{\mathrm{loc}}(M^d)$, and Y a P-special semimartingale with $[N^Y, [S]\,] \in \mathcal{A}_{\mathrm{loc}}(P)$. As for (4.1), we restrict our attention to solutions with $\psi^d \in L^2_{\mathrm{loc}}(M^d)$ and $L' \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$.

Proposition 4.1. *The BSDEs (4.1) and (4.2) are equivalent. More precisely, (Y, ψ, L) with $\psi \in L^2_{\mathrm{loc}}(M)$ and $L \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ solves (4.1) if and only if (Y, ψ^c, ψ^d, L') with $\psi^c \in L^2_{\mathrm{loc}}(M^c)$, $\psi^d \in L^2_{\mathrm{loc}}(M^d)$ and $L' \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ solves (4.2), where the tuples are related by*

$$\psi \cdot M + L = \psi^c \cdot M^c + \psi^d \cdot M^d + L'. \tag{4.3}$$

Proof. If (Y, ψ, L) solves (4.1), we use the Galtchouk–Kunita–Watanabe decomposition of $\psi \cdot M + L$ with respect to M^c and M^d to obtain (4.3) and define ψ^c, ψ^d, L'; so L' is strongly P-orthogonal to both M^c and M^d, and taking the covariation with M and using $\langle L, M\rangle \equiv 0$ gives $\psi = \psi^c\delta^c + \psi^d\delta^d$. Plugging this and (4.3) into (4.1) shows directly that (Y, ψ^c, ψ^d, L') solves (4.2).

Conversely, if (Y, ψ^c, ψ^d, L') solves (4.2), we define $\psi := \psi^c\delta^c + \psi^d(1-\delta^c) \in L^2_{\mathrm{loc}}(M)$ and $L := \psi^c \cdot M^c + \psi^d \cdot M^d + L' - \psi \cdot M \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$. Then plugging into (4.2) directly shows that (Y, ψ, L) satisfies (4.1), and since $\langle L, M\rangle \equiv 0$ due to the definitions above, L is also strongly P-orthogonal to M. So (Y, ψ, L) solves (4.1). **q.e.d.**

Equation (4.2) is particularly convenient for models with simple jumps, as illustrated by

Example 4.2. Consider the *jump-diffusion model*

$$dS_t = S_{t-}(\mu_t\,dt + \sigma_t\,dW_t + \eta_t\,dn_t), \qquad S_0 > 0,$$

where W is a Brownian motion and $n_t = N_t - \alpha t$, $0 \le t \le T$, is the compensated martingale of a simple Poisson process with intensity $\alpha > 0$. The predictable processes μ, σ, η satisfy $\sigma \neq 0$ and suitable integrability conditions, and we assume that $\eta > -1$ to ensure that $S > 0$. Then we have $dM^c_t = S_{t-}\sigma_t\,dW_t$, $dM^d_t = S_{t-}\eta_t\,dn_t$, $d\langle M\rangle_t = S^2_{t-}(\sigma^2_t + \alpha\eta^2_t)\,dt$, $\lambda_t = \frac{\mu_t}{S_{t-}(\sigma^2_t+\alpha\eta^2_t)}$ and $\delta^c_t = \frac{\sigma^2_t}{\sigma^2_t+\alpha\eta^2_t}$. Because $\langle M\rangle$ is continuous, so is B^Y due to (4.2); hence ${}^{\mathbf{P}}Y = Y_-$ by (2.10). Moreover, using $[n] = N$ gives

$$\begin{aligned}
\big[N^Y, [S]\big]^{\mathbf{P}}_t &= \big[\psi^c\cdot M^c + \psi^d\cdot M^d + L', [M^d]\big]^{\mathbf{P}}_t \\
&= \big[\psi^d\cdot M^d + L', (S_-\eta)^2\cdot[n]\big]^{\mathbf{P}}_t \\
&= (S^3_-\psi^d\eta^3)\cdot N^{\mathbf{P}}_t \\
&= (S^3_-\psi^d\eta^3\alpha)\cdot t
\end{aligned}$$

so that $g_t(Y) = \frac{\alpha\eta^3_t\psi^d_t S_{t-}}{\sigma^2_t+\alpha\eta^2_t}$. Using the notations $\tilde\psi^c = \psi^c S_-\sigma$, $\tilde\psi^d = \psi^d S_-\eta$ and plugging in then allows us to rewrite the BSDE (4.2) after simple calculations as

$$Y_t = Y_0 + \int_0^t \frac{(\tilde\psi^c_s\sigma_s + \alpha\tilde\psi^d_s\eta_s + \mu_s Y_{s-})^2}{Y_{s-}(\sigma^2_s + \alpha\eta^2_s) + \alpha\tilde\psi^d_s\eta^2_s}\,ds + \int_0^t \tilde\psi^c_s\,dW_s + \int_0^t \tilde\psi^d_s\,dn_s + L'_t, \qquad Y_T = 1.$$

It depends on the choice of the filtration $I\!F$ whether we can have a nontrivial $L' \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to both M^c and M^d, or W and n. If $I\!F$ is generated by W and N, then $L' \equiv 0$ automatically by the martingale representation theorem in $I\!F^{W,N}$. □

4.2. Using random measures

For models with more general jumps, the version (4.2) of the basic BSDE (4.1) is less useful because one cannot easily express $g(Y)$ in terms of integrands like in the preceding example. We therefore use semimartingale characteristics and in particular work with the jump measure of S. For the required notations and results, we refer to Chapter II of Jacod/Shiryaev (2003). We take $E = I\!R$ there so that $\widetilde\Omega = \Omega \times [0,T] \times I\!R$ with the σ-field $\widetilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}(I\!R)$, where $\mathcal{P}$ is the predictable σ-field on $\Omega \times [0,T]$.

Denote by μ^S the random measure associated with the jumps of S and by ν its P-compensator. Using Proposition II.2.9 of Jacod/Shiryaev (2003), we have

$$\nu(\omega, dt, dx) = F_t(\omega, dx)\,dB_t(\omega)$$

for a predictable increasing B null at 0. Moreover, (2.7) gives $\Delta S = \Delta M + \lambda\Delta\langle M\rangle$ and $(x^2\wedge 1)*\mu^S \ll [M]+\langle M\rangle$, and combining this with the construction of B in Jacod/Shiryaev (2003) and (2.6), we see that $B \ll \langle M\rangle$. We introduce the predictable processes

$$b := \frac{dB}{d\langle M\rangle}, \qquad \delta^c := \frac{d\langle M^c\rangle}{d\langle M\rangle}$$

and note that $[M^d] = \sum(\Delta M)^2 = (x-\lambda\Delta\langle M\rangle)^2 * \mu^S$ implies that

$$\langle M^d\rangle = (x-\lambda\Delta\langle M\rangle)^2 * \nu = \big(\int (x-\lambda\Delta\langle M\rangle)^2 F(dx)\big)\boldsymbol{\cdot} B,$$

so that $\langle M\rangle = \langle M^c\rangle + \langle M^d\rangle$ can be reformulated as

$$\delta^c_t + b_t \int (x-\lambda_t\Delta\langle M\rangle_t)^2 F_t(dx) = 1 \qquad P\otimes\langle M\rangle\text{-a.e.} \tag{4.4}$$

With the notation $\widehat{W}_t = \int_{I\!R} W_t(x)\nu(\{t\},dx)$, we now consider the backward equation

$$Y_t = Y_0 + \int_0^t \frac{\big(\varphi_s\delta^c_s + b_s\int x(W_s(x)-\widehat{W}_s)F_s(dx) + \lambda_s{}^{\mathbf{p}}Y_s\big)^2}{{}^{\mathbf{p}}Y_s\delta^c_s + b_s\int x^2\big({}^{\mathbf{p}}Y_s + W_s(x)-\widehat{W}_s\big)F_s(dx)}\,d\langle M\rangle_s \tag{4.5}$$

$$+\int_0^t \varphi_s\,dM^c_s + W*(\mu^S-\nu)_t + L'_t, \qquad Y_T = 1.$$

A *solution* of (4.5) is a priori a tuple (Y,φ,W,L') with $\varphi\in L^2_{\mathrm{loc}}(M^c)$, $W\in\mathcal{G}^1_{\mathrm{loc}}(\mu^S)$ (see (3.62) in Jacod (1979)), $L'\in\mathcal{M}^2_{0,\mathrm{loc}}(P)$ strongly P-orthogonal to M^c and to the space of stochastic integrals $\{\bar{W}*(\mu^S-\nu)\,|\,\bar{W}\in\mathcal{G}^2_{\mathrm{loc}}(\mu^S)\}$, and Y a P-special semimartingale with $[N^Y,[S]]\in\mathcal{A}_{\mathrm{loc}}(P)$. As before for (4.1) and (4.2), we restrict our attention to solutions with $W\in\mathcal{G}^2_{\mathrm{loc}}(\mu^S)$ and $L'\in\mathcal{M}^2_{0,\mathrm{loc}}(P)$.

In view of the next result, (4.5) seems the natural form of the BSDE (4.1) or (2.18) in the general case, because its generator is expressed in terms of integrands. Nevertheless, as seen in Section 2, the form (2.18) is more convenient for proving results via stochastic calculus.

Proposition 4.3. *The BSDEs (4.1) and (4.5) are equivalent. More precisely, (Y,ψ,L) with $\psi\in L^2_{\mathrm{loc}}(M)$ and $L\in\mathcal{M}^2_{0,\mathrm{loc}}(P)$ solves (4.1) if and only if (Y,φ,W,L') with $\varphi\in L^2_{\mathrm{loc}}(M^c)$, $W\in\mathcal{G}^2_{\mathrm{loc}}(\mu^S)$ and $L'\in\mathcal{M}^2_{0,\mathrm{loc}}(P)$ solves (4.5), where the tuples are related by*

$$\psi\boldsymbol{\cdot} M + L = \varphi\boldsymbol{\cdot} M^c + W*(\mu^S-\nu) + L'.$$

Proof. If (Y,ψ,L) solves (4.1), we take its martingale part $\psi\boldsymbol{\cdot} M+L$ and represent this as

$$\psi\boldsymbol{\cdot} M + L = \varphi\boldsymbol{\cdot} M^c + W*(\mu^S-\nu) + U*\mu^S + \widetilde{L} \tag{4.6}$$

with $\varphi \in L^2_{\mathrm{loc}}(M^c)$, $W \in \mathcal{G}^2_{\mathrm{loc}}(\mu^S)$, $U \in \mathcal{H}^2_{\mathrm{loc}}(\mu^S)$ (see Jacod (1979), §3.3b, pp. 101–102) and $\widetilde{L} \in \mathcal{M}^2_{0,\mathrm{loc}}(P)$ with $[\widetilde{L}, S] \equiv 0$. This is the so-called *Jacod decomposition*; see Jacod (1979), Théorème 3.75, or Theorem 2.4 in Choulli/Schweizer (2011) for a more detailed exposition.

We next express $g(Y)$ in terms of W and ν. Using (4.1) and (4.6) yields

$$\Delta N^Y_t = W_t(\Delta S_t) I_{\{\Delta S_t \neq 0\}} - \widehat{W}_t + U_t(\Delta S_t) I_{\{\Delta S_t \neq 0\}} + \Delta \widetilde{L}_t. \tag{4.7}$$

Moreover, $\sum \Delta \widetilde{L} (\Delta S)^2 = \Delta S \boldsymbol{\cdot} [\widetilde{L}, S] \equiv 0$ so that we get

$$\big[N^Y, [S]\big] = \sum \Delta N^Y (\Delta S)^2 = \big(x^2 (W(x) - \widehat{W})\big) * \mu^S + \big(x^2 U(x)\big) * \mu^S.$$

Because $[N^Y, [S]\,]$ is in $\mathcal{A}_{\mathrm{loc}}(P)$, this implies that $x^2 U(x)$ is in $\mathcal{H}^1_{\mathrm{loc}}(\mu^S)$ so that $(x^2 U(x)) * \mu^S$ is a local P-martingale by Jacod (1979), (3.73). Hence we obtain

$$\big[N^Y, [S]\big]^{\mathbf{p}} = \Big(\big(x^2 (W(x) - \widehat{W})\big) * \mu^S\Big)^{\mathbf{p}} = \big(x^2 (W(x) - \widehat{W})\big) * \nu = \Big(\int x^2 (W(x) - \widehat{W}) F(dx)\Big) \boldsymbol{\cdot} B,$$

and so $g_t(Y) = b_t \int x^2 (W_t(x) - \widehat{W}_t) F_t(dx)$. Moreover, $[S] = [S]^c + \sum (\Delta S)^2 = \langle M^c \rangle + x^2 * \mu^S$ gives $[S]^{\mathbf{p}} = \langle M^c \rangle + x^2 * \nu = (\delta^c + \int x^2 F(dx)\, b) \boldsymbol{\cdot} \langle M \rangle$ so that comparing with (2.8) yields that $1 + \lambda^2 \Delta \langle M \rangle = \delta^c + b \int x^2 F(dx)$ and hence

$$\mathcal{N}_t(Y) = {}^{\mathbf{p}}Y_t (1 + \lambda_t^2 \Delta \langle M \rangle_t) + g_t(Y) = {}^{\mathbf{p}}Y_t \delta^c_t + b_t \int x^2 \big({}^{\mathbf{p}}Y_t + W_t(x) - \widehat{W}_t\big) F_t(dx). \tag{4.8}$$

If we now define $L' := U * \mu^S + \widetilde{L}$, then (4.6) gives

$$\psi \boldsymbol{\cdot} M + L = \varphi \boldsymbol{\cdot} M^c + W * (\mu^S - \nu) + L'. \tag{4.9}$$

But $[L', M] = [L', S] - [L', \lambda \boldsymbol{\cdot} \langle M \rangle] = (x U(x)) * \mu^S + [\widetilde{L}, S] - [L', \lambda \boldsymbol{\cdot} \langle M \rangle]$ is a local P-martingale by Yoeurp's lemma and a similar argument as just above, using now that $U \in \mathcal{H}^2_{\mathrm{loc}}(\mu^S)$; so $\langle L', M \rangle \equiv 0$ and L' is strongly P-orthogonal to M^c. Moreover, we have for all $\bar{W} \in \mathcal{G}^2_{\mathrm{loc}}(\mu^S)$ that $[\widetilde{L}, \bar{W} * (\mu^S - \nu)] = 0$ since $[\widetilde{L}, S] \equiv 0$, and so $\langle L', \bar{W} * (\mu^S - \nu) \rangle = \langle U * \mu^S, \bar{W} * (\mu^S - \nu) \rangle \equiv 0$ for all $\bar{W} \in \mathcal{G}^2_{\mathrm{loc}}(\mu^S)$ by Jacod (1979), Exercice 3.23. Finally, (2.7) and Yoeurp's lemma yield

$$\begin{aligned} \langle W * (\mu^S - \nu), M \rangle &= [W * (\mu^S - \nu), S - \lambda \boldsymbol{\cdot} \langle M \rangle]^{\mathbf{p}} \\ &= [W * (\mu^S - \nu), S]^{\mathbf{p}} \\ &= \Big(\big(x (W(x) - \widehat{W})\big) * \mu^S\Big)^{\mathbf{p}} \\ &= \big(x (W(x) - \widehat{W})\big) * \nu. \end{aligned} \tag{4.10}$$

Taking in (4.9) the covariation with M and using also $\langle L, M \rangle \equiv 0 \equiv \langle L', M \rangle$ yields

$$\psi \boldsymbol{\cdot} \langle M \rangle = \Big(\varphi \delta^c + \big(\int x (W(x) - \widehat{W}) F(dx)\big) b\Big) \boldsymbol{\cdot} \langle M \rangle$$

so that we get

$$\psi_t = \varphi_t \delta^c_t + b_t \int x(W_t(x) - \widehat{W}_t) F_t(dx) \qquad P \otimes \langle M \rangle\text{-a.e.} \tag{4.11}$$

Plugging (4.11) and (4.8) into (4.1) and using (4.9), we see that (Y, φ, W, L') solves (4.5).

Conversely, if (Y, φ, W, L') solves (4.5), then we define ψ by (4.11) and

$$L := \varphi \cdot M^c - \psi \cdot M + W * (\mu^S - \nu) + L'.$$

Then $\psi \in L^2_{\text{loc}}(M)$ due to (4.4) and because $W \in \mathcal{G}^2_{\text{loc}}(\mu^S)$, and so $L \in \mathcal{M}^2_{0,\text{loc}}(P)$. Moreover, (4.10), the definitions of L and ψ via (4.11) and the definitions of δ^c and b yield

$$\langle L, M \rangle = \langle L', M \rangle = \langle L', M^c + M^d \rangle = \langle L', M^c \rangle + \langle L', x * (\mu^S - \nu) \rangle \equiv 0$$

by the orthogonality properties of L', so that L is strongly P-orthogonal to M. Finally, the Jacod decomposition applied to L' implies that the latter must have the form $L' = U * \mu^S + \widetilde{L}$ due to its orthogonality properties. But then we obtain from (4.5) again (4.7), hence also (4.8), and then plugging in shows that (Y, ψ, L) solves (4.1). This completes the proof. **q.e.d.**

Just for completeness, but without any details, we give here the equivalent versions of the BSDEs (3.2) and (3.3) for $v^{(1)}$ and $v^{(0)}$. They are

$$\begin{aligned} dY^{(1)}_t = {} & \frac{\big(\varphi^{(1)}_t \delta^c_t + b_t \int x(W^{(1)}_t(x) - \widehat{W^{(1)}_t}) F_t(dx) + \lambda_t {}^{\mathbf{P}}Y^{(1)}_t\big)}{{}^{\mathbf{P}}Y^{(2)}_t \delta^c_t + b_t \int x^2 \big({}^{\mathbf{P}}Y^{(2)}_t + W^{(2)}_t(x) - \widehat{W^{(2)}_t}\big) F_t(dx)} \\ & \times \big(\varphi^{(2)}_t \delta^c_t + b_t \int x(W^{(2)}_t(x) - \widehat{W^{(2)}_t}) F_t(dx) + \lambda_t {}^{\mathbf{P}}Y^{(2)}_t\big)\, d\langle M \rangle_t \\ & + \varphi^{(1)}_t\, dM^c_t + d\big(W^{(1)} * (\mu^S - \nu)\big)_t + dL^{(1),\prime}_t, \quad Y^{(1)}_T = H, \end{aligned}$$

and

$$dY^{(0)}_t = \frac{\big(\varphi^{(1)}_t \delta^c_t + b_t \int x(W^{(1)}_t(x) - \widehat{W^{(1)}_t}) F_t(dx) + \lambda_t {}^{\mathbf{P}}Y^{(1)}_t\big)^2}{{}^{\mathbf{P}}Y^{(2)}_t \delta^c_t + b_t \int x^2 \big({}^{\mathbf{P}}Y^{(2)}_t + W^{(2)}_t(x) - \widehat{W^{(2)}_t}\big) F_t(dx)}\, d\langle M \rangle_t + dN^{(0)}_t, \quad Y^{(0)}_T = H^2.$$

Finally, the recursive representation for the optimal strategy in (3.6) takes the form

$$\begin{aligned} \vartheta^{*,0}_t = {} & \frac{\varphi^{(1)}_t \delta^c_t + b_t \int x(W^{(1)}_t(x) - \widehat{W^{(1)}_t}) F_t(dx) + \lambda_t {}^{\mathbf{P}}Y^{(1)}_t}{{}^{\mathbf{P}}Y^{(2)}_t \delta^c_t + b_t \int x^2 \big({}^{\mathbf{P}}Y^{(2)}_t + W^{(2)}_t(x) - \widehat{W^{(2)}_t}\big) F_t(dx)} \\ & - \frac{\varphi^{(2)}_t \delta^c_t + b_t \int x(W^{(2)}_t(x) - \widehat{W^{(2)}_t}) F_t(dx) + \lambda_t {}^{\mathbf{P}}Y^{(2)}_t}{{}^{\mathbf{P}}Y^{(2)}_t \delta^c_t + b_t \int x^2 \big({}^{\mathbf{P}}Y^{(2)}_t + W^{(2)}_t(x) - \widehat{W^{(2)}_t}\big) F_t(dx)} X^{\vartheta^{*,0}}_{t-}. \end{aligned}$$

Of course, this can equivalently be rewritten as a linear SDE for $X^{\vartheta^{*,0}}$ as in (3.4), simply by integrating with respect to S.

4.3. Further comments

At this point, it seems appropriate to comment on related work in the literature, where we restrict ourselves to papers that have used BSDE techniques in the context of mean-variance hedging. While extending work by many authors done for an Itô process setting in a Brownian filtration, the results in Mania/Tevzadze (2003a,b) and Bobrovnytska/Schweizer (2004) still all assume that S is continuous. At the other end of the scale, Černý/Kallsen (2007) have a general $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$, with $I\!\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$; but their methods do not exploit stochastic control ideas and results at all, and BSDEs appear only very tangentially in their equations (3.32) and (3.37). As a matter of fact, their opportunity process L equals our coefficient $v^{(2)}$ and so their equation (3.37), which gives a BSDE for L, should coincide with our equation (4.5). However, Černý/Kallsen (2007) give no proof for (3.37) and even remark that "it is not obvious whether this representation is of any use". Moreover, a closer examination shows that (3.37) is not entirely correct; it seems that they dropped the jumps of the FV part of L somewhere, which explains why their equation has L_- instead of (the correct term) ${}^{\mathbf{p}}L$.

The paper closest to our work is probably Kohlmann/Xiong/Ye (2010). They first study the variance-optimal martingale measure as in Mania/Tevzadze (2003b) via the problem dual to mean-variance hedging and obtain a BSDE that describes $\widetilde{V} = 1/V^0(1) = 1/v^{(2)}$; see our Proposition 2.2. For mean-variance hedging itself, they subsequently describe the optimal strategy in feedback form with the help of a process (called h) for which they give a BSDE. Their assumptions are considerably more restrictive than ours because in addition to $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $I\!\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$, they also suppose that S is quasi-left-continuous; and for the results on mean-variance hedging, they additionally even assume that $\mathcal{M}^d_{\mathrm{loc}}(P)$ is generated by integrals of $\mu^S - \nu$ (and also that the VOMM exists and satisfies the reverse Hölder inequality $R_2(P)$ and a certain jump condition). We found it hard to see exactly why this restrictive condition on $\mathcal{M}^d_{\mathrm{loc}}(P)$ is needed; the proof in Kohlmann/Xiong/Ye (2010) for their verification result is rather computational and does not explain where the rather technical BSDEs come from.

Finally, a similar (subjective) comment as the last one also applies to Lim (2005). The problem studied there is mean-variance hedging (not the VOMM), and the process S is a multivariate version of the simple jump-diffusion model in Example 4.2, with a d-dimensional Brownian motion W and an m-variate Poisson process N. The filtration used for strategies ϑ and payoffs H is generated by W and N; but all model coefficients (including the intensity of N) are assumed to be $I\!\!F^W$-predictable. Technically speaking, this condition serves to simplify Lim's equation (3.1), which corresponds to our equation from Example 4.2 for Y without the jump term. It would be interesting to see also at the conceptual level why the assumption is needed.

Remark. As already pointed out before Theorem 3.1, the BSDE system (3.1)–(3.3) is less complicated than it looks. It is only weakly coupled, meaning that one can solve (3.3) (even directly) once one has the solutions of (3.1) and (3.2), and that (3.2) is linear and hence also readily solved once one has the solution of (3.1). In general, however, (3.1) has a very complicated driver, and it seems a genuine challenge for abstract BSDE theory to prove existence of a solution directly via BSDE techniques. We do not do that (and do not need to) since we only use the BSDEs to describe optimal strategies; existence of the latter (and hence existence of solutions to the BSDEs) is proved directly via other arguments.

In the special case where the filtration $I\!\!F$ is continuous, the complicated equation (3.1) or (2.18) can be reduced to a classical quadratic BSDE, as follows. First of all, as already pointed out before Lemma 2.3, the operation $\mathcal{N}(Y)$ in (2.12) reduces to $\mathcal{N}(Y) = Y$, at least in the context of (2.18). So (2.18) becomes

$$dY_t = \frac{(\psi_t + \lambda_t Y_t)^2}{Y_t}\, d\langle M\rangle_t + \psi_t\, dM_t + dL_t, \qquad Y_T = 1, \tag{4.12}$$

and we know from Lemma 2.1 that the solution $q = V^0(1)$ is strictly positive. If we introduce $y := \log Y$, apply Itô's formula and define $\varphi := \psi/Y$, $\ell := \int (1/Y)\, dL$, then it is straightforward to verify that (4.12) can be rewritten as

$$dy_t = \varphi_t\, dM_t + \big((\varphi_t + \lambda_t)^2 - \tfrac{1}{2}\varphi_t^2\big)\, d\langle M\rangle_t + d\ell_t - \tfrac{1}{2}\, d\langle \ell\rangle_t, \qquad y_T = 0.$$

This can then be tackled by standard BSDE methods if desired. ◇

5. Examples

In this section, we present some simple examples and special cases to illustrate our results. We keep this deliberately short in view of the total length of the paper. **Throughout this section, we assume that $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $I\!\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$.**

Recall the P-canonical decomposition $S = S_0 + M + \int \lambda\, d\langle M\rangle$ of our price process. Because $\lambda \in L^2_{\mathrm{loc}}(M)$, the process $\widehat{Z} := \mathcal{E}(-\lambda \cdot M)$ is in $\mathcal{M}^2_{\mathrm{loc}}(P)$ with $\widehat{Z}_0 = 1$. Moreover, it is easy to check that $\widehat{Z}S$ is a local P-martingale so that $\widehat{Z}$ is a so-called signed local martingale density for S. If $\widehat{Z}$ is a true P-martingale and in $\mathcal{M}^2(P)$, then $\widehat{Q}$ with $d\widehat{Q} := \widehat{Z}_T\, dP$ is in $I\!\!P^2_{\mathrm{s},\sigma}(S)$ and called the *minimal signed (local) martingale measure* for S; if even $\widehat{Z} > 0$ so that $\widehat{Q}$ is in $I\!\!P^2_{\mathrm{e},\sigma}(S)$, then $\widehat{Q}$ is the *minimal martingale measure (MMM)* for S.

The MMM is very convenient because its density process $\widehat{Z}$ can be read off explicitly from S. On the other hand, the important quantity for mean-variance hedging is the variance-optimal martingale measure (VOMM) $\widetilde{Q}$. By Proposition 2.6, we could construct a solution to the BSDE (2.18) from $\widetilde{Q}$ by

$$V^0_t(1) = q_t = v^{(2)}_t = 1/\widetilde{V}_t = \frac{(Z^{\widetilde{Q}}_t)^2}{E\big[(Z^{\widetilde{Q}}_T)^2 \,\big|\, \mathcal{F}_t\big]}, \qquad 0 \le t \le T,$$

but the density process $Z^{\widetilde{Q}}$ is usually difficult to find. An exception is the case when $\widetilde{Q} = \widehat{Q}$, since then $Z^{\widetilde{Q}} = \widehat{Z} = \mathcal{E}(-\lambda \cdot M)$ and the above formula allows us to find an explicit expression for $v^{(2)}$. To make this approach work, we need conditions when $\widetilde{Q}$ and $\widehat{Q}$ coincide. This has been studied before, and we could give some new results, but do not do so here for reasons of space. We only mention the MMM since it comes up later in another example.

5.1. Easy solutions for the process $V^0(1) = v^{(2)}$

In terms of complexity, the BSDE (2.18) or one of its equivalent forms (3.1), (4.2), (4.5) is the most difficult one. So we focus on that equation, in the form (4.5), and we try to have a solution tuple (Y, φ, W, L') with $\varphi \equiv 0$ and $W \equiv 0$. Then (4.5) simplifies to

$$Y_t = Y_0 + \int_0^t \frac{\lambda_s^2\, {}^{\mathbf{p}}Y_s}{1+\lambda_s^2 \Delta\langle M\rangle_s}\, d\langle M\rangle_s + L'_t,$$

which gives $\Delta B^Y = \frac{\lambda^2\, {}^{\mathbf{p}}Y}{1+\lambda^2\Delta\langle M\rangle}\Delta\langle M\rangle$. But ${}^{\mathbf{p}}Y = Y_- + \Delta B^Y$ by (2.10), and plugging this in above and solving for ΔB^Y allows us to get ${}^{\mathbf{p}}Y = Y_-(1+\lambda^2\Delta\langle M\rangle)$ so that (4.5) becomes

$$Y_t = Y_0 + \int_0^t Y_{s-}\lambda_s^2\, d\langle M\rangle_s + L'_t, \qquad Y_T = 1. \tag{5.1}$$

This is the equation for a generalised stochastic exponential, and so it is not surprising that we can find an explicit solution.

Corollary 5.1. *Set $K := \langle \lambda \cdot M\rangle$ and suppose that*

$$\mathcal{E}(K)_T^{-1} = c + m_T$$

with a constant $c > 0$ and a P-martingale m which is strongly P-orthogonal both to M^c and to the space of stochastic integrals $\{\bar{W} * (\mu^S - \nu) \,|\, \bar{W} \in \mathcal{G}^2_{\mathrm{loc}}(\mu^S)\}$. Then the solution of (4.5) is given by $\varphi \equiv 0$, $W \equiv 0$ and

$$Y_t = E[\mathcal{E}(K)_t/\mathcal{E}(K)_T \,|\, \mathcal{F}_t] = \mathcal{E}(K)_t(c+m_t), \quad L'_t = \int_0^t \mathcal{E}(K)_{s-}\, dm_s + [\mathcal{E}(K), m]_t. \tag{5.2}$$

Proof. Since (5.1) can be written as $Y = Y_0 + \int Y_-\, dK + L'$, defining Y and L' by (5.2) gives by the product rule that (Y, L') satisfy (5.1) with $Y_T = 1$, and L' is a local P-martingale like m by Yoeurp's lemma. Finally, for every $\bar{W} \in \mathcal{G}^2_{\mathrm{loc}}(\mu^S)$, we have that

$$\big[\bar{W} * (\mu^S - \nu), [\mathcal{E}(K), m]\big] = \sum \Delta\big(\bar{W} * (\mu^S - \nu)\big)\Delta\mathcal{E}(K)\Delta m = \Delta\mathcal{E}(K) \cdot [\bar{W} * (\mu^S - \nu), m]$$

is a local P-martingale because m is strongly P-orthogonal to $\bar{W} * (\mu^S - \nu)$. Hence L' is also strongly P-orthogonal to $\bar{W} * (\mu^S - \nu)$, and so $(Y, 0, 0, L')$ is a solution to (4.5). **q.e.d.**

Example 5.2. A special case of Corollary 5.1 occurs if the (final) *mean-variance tradeoff* $\langle\lambda\cdot M\rangle_T$ and all the jumps $\lambda^2\Delta\langle M\rangle$ are *deterministic*. Then $m\equiv 0$, the solution for Y is

$$Y_t=\mathcal{E}(\langle\lambda\cdot M\rangle)_t/\mathcal{E}(\langle\lambda\cdot M\rangle)_T,\qquad 0\le t\le T$$

(which is adapted because $\mathcal{E}(\langle\lambda\cdot M\rangle)_T$ is deterministic), and all other quantities in the BSDEs (2.18) or (4.2) or (4.5) are identically 0. If S or M or even only $A=\int\lambda^2\,d\langle M\rangle$ is continuous, the above expression simplifies to

$$Y_t=e^{\langle\lambda\cdot M\rangle_t-\langle\lambda\cdot M\rangle_T},\qquad 0\le t\le T.$$

□

Similar results as in this section, but under more restrictive assumptions, have been obtained by several authors. We only mention exemplarily the work of Biagini/Guasoni/Pratelli (2000), Mania/Tevzadze (2003b) and Santacroce (2006).

5.2. The discrete-time case

Now we briefly look at the special case of a model in finite discrete time $k=0,1,\dots,T$. Our price process is given by $S=(S_k)_{k=0,1,\dots,T}$, and we assume as in (2.7) that

$$S=S_0+M+\lambda\cdot\langle M\rangle \tag{5.3}$$

with a martingale $M=(M_k)_{k=0,1,\dots,T}$ null at 0. We assume that S is square-integrable to avoid technical complications, and we write $\Delta_k Y:=Y_k-Y_{k-1}$ for the increments of a process $Y=(Y_k)_{k=0,1,\dots,T}$. The Doob decomposition $S=S_0+M+A$ is then given by $\Delta_k A=E[\Delta_k S\,|\,\mathcal{F}_{k-1}]$, we have $\Delta_k\langle M\rangle=E[(\Delta_k M)^2\,|\,\mathcal{F}_{k-1}]=\mathrm{Var}[\Delta_k S\,|\,\mathcal{F}_{k-1}]$, and so (5.3) takes the form $S=S_0+M+\sum\limits_j\lambda_j\Delta_j\langle M\rangle$ with

$$\lambda_j=\frac{\Delta_j A}{\Delta_j\langle M\rangle}=\frac{E[\Delta_j S\,|\,\mathcal{F}_{j-1}]}{\mathrm{Var}[\Delta_j S\,|\,\mathcal{F}_{j-1}]}. \tag{5.4}$$

For the discrete-time version of the BSDE (2.18), we need ${}^{\mathrm{p}}Y_j=E[Y_j\,|\,\mathcal{F}_{j-1}]$ and the density $g(Y)$ of $[N^Y,[S]\,]^{\mathrm{p}}$ with respect to $\langle M\rangle$. But $[N^Y,[S]\,]=\sum\limits_j(\Delta_j N^Y)(\Delta_j S)^2$ so that

$$g_j(Y)\Delta_j\langle M\rangle=E[(\Delta_j N^Y)(\Delta_j S)^2\,|\,\mathcal{F}_{j-1}]. \tag{5.5}$$

Moreover, we have

$$(1+\lambda_j^2\Delta_j\langle M\rangle)\Delta_j\langle M\rangle=\mathrm{Var}[\Delta_j S\,|\,\mathcal{F}_{j-1}]+(E[\Delta_j S\,|\,\mathcal{F}_{j-1}])^2=E[(\Delta_j S)^2\,|\,\mathcal{F}_{j-1}], \tag{5.6}$$

and the Galtchouk–Kunita–Watanabe decomposition $N^Y = \sum_j \psi_j \Delta_j M + L$ yields

(5.7) $\psi_j \Delta_j \langle M \rangle = \mathrm{Cov}(\Delta_j N^Y, \Delta_j M \,|\, \mathcal{F}_{j-1}) = \mathrm{Cov}(\Delta_j Y, \Delta_j S \,|\, \mathcal{F}_{j-1}) = \mathrm{Cov}(Y_j, \Delta_j S \,|\, \mathcal{F}_{j-1})$.

Hence we get

$$\begin{aligned}(\psi_j + \lambda_j{}^{\mathrm{P}}Y_j)^2 (\Delta_j \langle M \rangle)^2 &= \big(\mathrm{Cov}(Y_j, \Delta_j S \,|\, \mathcal{F}_{j-1}) + E[\Delta_j S \,|\, \mathcal{F}_{j-1}] E[Y_j \,|\, \mathcal{F}_{j-1}]\big)^2 \\ &= (E[Y_j \Delta_j S \,|\, \mathcal{F}_{j-1}])^2.\end{aligned}$$

Writing out the discrete-time analogue of (2.18), expanding the ratios in the first appearing sum with $\Delta_j \langle M \rangle$ and using (5.4)–(5.7) then yields

$$\begin{aligned}(5.8) \qquad Y_k &= Y_0 + \sum_{j=1}^k \frac{(\psi_j + \lambda_j{}^{\mathrm{P}}Y_j)^2}{{}^{\mathrm{P}}Y_j(1 + \lambda_j^2 \Delta_j \langle M \rangle) + g_j(Y)} \Delta_j \langle M \rangle + \sum_{j=1}^k \psi_j \Delta_j M + L_k \\ &= Y_0 + \sum_{j=1}^k \frac{(E[Y_j \Delta_j S \,|\, \mathcal{F}_{j-1}])^2}{E[Y_j \,|\, \mathcal{F}_{j-1}] E[(\Delta_j S)^2 \,|\, \mathcal{F}_{j-1}] + E[(\Delta_j N^Y)(\Delta_j S)^2 \,|\, \mathcal{F}_{j-1}]} \\ &\quad + \sum_{j=1}^k \psi_j \Delta_j M + L_k, \qquad Y_T = 1.\end{aligned}$$

But $Y_j = Y_{j-1} + \Delta_j N^Y + \Delta_j B^Y$ gives $E[Y_j \,|\, \mathcal{F}_{j-1}] = Y_{j-1} + \Delta_j B^Y = N^Y_{j-1} + B^Y_j$, and the denominator in the third sum in (5.8) therefore equals

$$E[(N^Y_{j-1} + B^Y_j + \Delta_j N^Y)(\Delta_j S)^2 \,|\, \mathcal{F}_{j-1}] = E[Y_j (\Delta_j S)^2 \,|\, \mathcal{F}_{j-1}].$$

Passing to increments and taking conditional expectations to make the martingale increments vanish, the equation (5.8) thus can be written as

$$Y_{k-1} = E[Y_k - \Delta_k Y \,|\, \mathcal{F}_{k-1}] = E[Y_k \,|\, \mathcal{F}_{k-1}] - \frac{(E[Y_k \Delta_k S \,|\, \mathcal{F}_{k-1}])^2}{E[Y_k (\Delta_k S)^2 \,|\, \mathcal{F}_{k-1}]}, \qquad Y_T = 1.$$

This is exactly the recursive relation derived in equation (3.1) in Theorem 1 of Gugushvili (2003); see also equation (3.36) in Černý/Kallsen (2007). Under more restrictive assumptions, analogous equations have also been obtained in equation (5) in Theorem 2 of Černý (2004) or in equation (2.19) in Theorem 1 of Bertsimas/Kogan/Lo (2001).

5.3. On the relation to Arai (2005)

Our final example serves to illustrate the relations between our work and that of Arai (2005), whose assumptions are rather similar to ours. More precisely, Arai (2005) assumes that S (which he calls X) is locally bounded, and that the VOMM $\widetilde{Q}$ exists in $I\!P^2_{\mathrm{e},\sigma}(S)$ and satisfies the reverse Hölder inequality $R_2(P)$ and a condition on the jumps of $Z^{\widetilde{Q}}$. This implies of course $S \in \mathcal{S}^2_{\mathrm{loc}}(P)$ and $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$. Arai (2005) does not use BSDEs, but works with a change of numeraire as in Gouriéroux/Laurent/Pham (1998). His numeraire is $E_{\widetilde{Q}}[Z^{\widetilde{Q}}_T \,|\, \mathcal{F}_.]$, and to ensure that this is positive, the existence of the VOMM $\widetilde{Q}$ in $I\!P^2_{\mathrm{e},\sigma}(S)$ is needed. The example below illustrates that our assumptions are strictly weaker than those of Arai (2005).

Example 5.3. We start with two independent simple Poisson processes $N^{(\pm)}$ with the same intensity $\alpha > 0$ and define $n^\pm_t := N^{(\pm)}_t - \alpha t$, $0 \le t \le T$. We then set

$$dS_t = S_{t-}(\gamma_+ \, dn^+_t - \gamma_- \, dn^-_t + \delta \, dt) =: S_{t-} \, dR_t,$$

so that S is clearly locally bounded, hence in $\mathcal{S}^2_{\mathrm{loc}}(P)$, and even quasi-left-continuous. We claim that we can choose the parameters $\alpha, \gamma_+, \gamma_-, \delta$ such that

1) $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$,
2) the variance-optimal signed martingale measure $\widetilde{Q} \in I\!P^2_{\mathrm{s},\sigma}(S)$ coincides with the minimal signed martingale measure $\widehat{Q}$, but is not in $I\!P^2_{\mathrm{e},\sigma}(S)$, which means in our terminology and that of Arai (2005) that the VOMM does not exist.

Let us first argue 2). Since $dM_t = S_{t-}(\gamma_+ \, dn^+_t - \gamma_- \, dn^-_t)$ gives $d\langle M\rangle_t = S^2_{t-}(\gamma^2_+ + \gamma^2_-)\alpha \, dt$ and we have $dA_t = S_{t-}\delta \, dt$, we obtain

$$\lambda \cdot M = \frac{\delta}{\alpha(\gamma^2_+ + \gamma^2_-)}(\gamma_+ n^+ - \gamma_- n^-).$$

So as soon as we have

$$\frac{\delta \gamma_+}{\alpha(\gamma^2_+ + \gamma^2_-)} > 1, \tag{5.9}$$

we get $-\lambda \Delta M < -1$ at jumps of $N^{(+)}$ so that $\widehat{Z} = \mathcal{E}(-\lambda \cdot M)$ also takes negative values. Because the mean-variance tradeoff process $\langle \lambda \cdot M\rangle_t = \frac{\delta^2}{\alpha(\gamma^2_+ + \gamma^2_-)} t$, $0 \le t \le T$, is deterministic, the signed MMM $\widehat{Q}$ is variance-optimal by Theorem 8 of Schweizer (1995). Moreover, $\widehat{Z}$ is clearly in $\mathcal{M}^2(P)$ and so $\widetilde{Q} = \widehat{Q}$ is in $I\!P^2_{\mathrm{s},\sigma}(S)$, but not in $I\!P^2_{\mathrm{e},\sigma}(S)$. This gives 2).

To construct an element of $I\!P^2_{\mathrm{e},\sigma}(S)$, we start with $Z := \mathcal{E}(L) := \mathcal{E}(\beta_1 n^+ + \beta_2 n^-)$, which is clearly in $\mathcal{M}^2(P)$. To ensure that $Z > 0$, we need $\beta_1 > -1$ and $\beta_2 > -1$. Next, the product ZS is by Itô's formula seen to be a local P-martingale if and only if $\delta \, dt + d\langle L, R\rangle_t \equiv 0$, which translates into the condition $\delta = (\beta_2 \gamma_- - \beta_1 \gamma_+)\alpha$. This allows us to rewrite (5.8) as

$$\frac{\gamma^2_+ + \gamma^2_-}{\gamma_+} < \frac{\delta}{\alpha} = \beta_2 \gamma_- - \beta_1 \gamma_+,$$

and if we choose $\gamma_+ = \gamma_- = \gamma$, this boils down to $\beta_2 - \beta_1 > 2$ and $\frac{\delta}{\alpha} = (\beta_2 - \beta_1)\gamma$. By the Bayes rule, S is then a local Q-martingale under $Q \approx P$ with $dQ = Z_T \, dP$.

If we now choose $\varepsilon > 0$ and $\beta_1 = \beta > -1$, $\beta_2 = \beta + 2 + \varepsilon$, $\alpha = 1$, $\delta = (2+\varepsilon)\gamma$, one readily verifies that all conditions above are satisfied; hence $I\!P^2_{\mathrm{e},\sigma}(S) \neq \emptyset$ since it contains Q. If we take $\gamma \in (0,1)$, we even keep $S > 0$ since $\Delta R > -1$. $\square$

Remark. By its construction, the minimal martingale density $\widehat{Z}$ is always based on $-\lambda \cdot M$. With our above choice of model parameters $\gamma_+ = \gamma_- = \gamma$, this is symmetric in n^+ and $-n^-$ and therefore risks getting negative jumps rather easily. In contrast, writing

$$L = \beta n^+ + (\beta + 2 + \varepsilon) n^- = -\lambda \cdot M + \widetilde{L}$$

with $\widetilde{L} = (\beta + 1 + \frac{\varepsilon}{2}) n^+ + (\beta + 1 + \frac{\varepsilon}{2}) n^-$ shows that it can be very beneficial to have some extra freedom when choosing an ELMM or a martingale density. This is quite analogous to the well-known counterexample in Delbaen/Schachermayer (1998). $\diamond$

Acknowledgements. Monique Jeanblanc and Michael Mania acknowledge financial support from Chaire de Crédit, French Banking Federation. Michael Mania also acknowledges the support from the Georgian National Foundation Grant N. ST09_471_3-104. Marina Santacroce is partially supported by MIUR, grant 2008YYYBE4. Martin Schweizer gratefully acknowledges financial support by the Swiss Finance Institute (SFI) and the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management). The NCCR FINRISK is a research instrument of the Swiss National Science Foundation.

References

T. Arai (2005), "An extension of mean-variance hedging to the discontinuous case", *Finance and Stochastics 9*, 129–139

D. Bertsimas, L. Kogan and A. W. Lo (2001), "Hedging derivative securities and incomplete markets: An ε-arbitrage approach", *Operations Research 49*, 372–397

F. Biagini, P. Guasoni and M. Pratelli (2000), "Mean-variance hedging for stochastic volatility models", *Mathematical Finance 10*, 109–123

O. Bobrovnytska and M. Schweizer (2004), "Mean-variance hedging and stochastic control: Beyond the Brownian setting", *IEEE Transactions on Automatic Control 49*, 396–408

A. Černý (2004), "Dynamic programming and mean-variance hedging in discrete time", *Applied Mathematical Finance 11*, 1–25

A. Černý and J. Kallsen (2007), "On the structure of general mean-variance hedging strategies", *Annals of Probability 35*, 1479–1531

T. Choulli, L. Krawczyk and C. Stricker (1998), "$\mathcal{E}$-martingales and their applications in mathematical finance", *Annals of Probability 26*, 853–876

T. Choulli and M. Schweizer (2011), "Stability of σ-martingale densities in $L \log L$ under an equivalent change of measure", *NCCR FINRISK working paper No. 676, ETH Zürich*, `http://www.nccr-finrisk.uzh.ch/media/pdf/wp/WP676_D1.pdf`

F. Delbaen and W. Schachermayer (1996), "The variance-optimal martingale measure for continuous processes", *Bernoulli 2*, 81–105; Amendments and corrections (1996), *Bernoulli 2*, 379–380

F. Delbaen and W. Schachermayer (1998), "A simple counterexample to several problems in the theory of asset pricing", *Mathematical Finance 8*, 1-11

C. Dellacherie and P.-A. Meyer (1978), "Probabilities and Potential", *North-Holland, Amsterdam*

C. Dellacherie and P.-A. Meyer (1982), "Probabilities and Potential B. Theory of Martingales", *North-Holland, Amsterdam*

N. El Karoui (1981), "Les aspects probabilistes du contrôle stochastique", *Ecole d'Eté de Probabilités de Saint-Flour IX-1979, Lecture Notes in Mathematics 876, Springer*, 73–238

C. Gouriéroux, J. P. Laurent and H. Pham (1998), "Mean-variance hedging and numéraire", *Mathematical Finance 8*, 179–200

S. Gugushvili (2003), "Dynamic programming and mean-variance hedging in discrete time", *Georgian Mathematical Journal 10*, 237–246

J. Jacod (1979), "Calcul Stochastique et Problèmes de Martingales", *Lecture Notes in Mathematics 714, Springer*

J. Jacod and A. N. Shiryaev (2003), "Limit Theorems for Stochastic Processes", *second edition, Springer*

M. Kohlmann, D. Xiong and Z. Ye (2010), "Mean variance hedging in a general jump model", *Applied Mathematical Finance 17*, 29–57

A. E. B. Lim (2005), "Mean-variance hedging when there are jumps", *SIAM Journal on Control and Optimization 44*, 1893–1922

M. Mania and R. Tevzadze (2003a), “Backward stochastic PDE and imperfect hedging”, *International Journal of Theoretical and Applied Finance 6*, 663–692

M. Mania and R. Tevzadze (2003b), “A semimartingale backward equation and the variance-optimal martingale measure under general information flow”, *SIAM Journal on Control and Optimization 42*, 1703-1726

M. Santacroce (2006), “Derivatives pricing via p-optimal martingale measures: some extreme cases”, *Journal of Applied Probability 43*, 634–651

M. Schweizer (1995), “On the minimal martingale measure and the Föllmer–Schweizer decomposition”, *Stochastic Analysis and Applications 13*, 573–599

M. Schweizer (2010), “Mean-variance hedging”, *in: R. Cont (ed.), “Encyclopedia of Quantitative Finance”, Wiley*, 1177–1181